

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2006, 10:00:16 ; Search time 900.103 Seconds

(without alignment)

Title: US-10-601-319-1

Perfect score: 1322
Sequence: 1 atgaagcattatcc.....atcacatcacatcaaa 1323

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_8:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2004as:*
- 12: geneseqn2004bs:*
- 13: geneseqn2004cs:*
- 14: geneseqn2005s:*
- 15: geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1322	100.0	1323	5 AAC88885	AAC88885 Escherich
2	1322	100.0	1323	6 AAD36473	AAD36473 Escherich
3	1322	100.0	1323	6 AAD25460	AAD25460 Escherich
4	1322	100.0	1323	8 ADA19445	ADA19445 E. coli B
5	1322	100.0	1323	12 ADO50291	ADO50291 Escherich
6	1322	100.0	1323	15 AEE75412	AEE75412 E. coli B
7	1315.6	99.5	1323	12 ADO50301	ADO50301 Kangaroo
8	1315.6	99.5	1323	12 ADO50303	ADO50303 Escherich
9	1287.8	97.4	1901	6 AAD25463	AAD25463 Escherich
10	1287.8	97.4	1901	6 ADA19449	ADA19449 E. coli K
11	1287.8	97.4	1901	10 ADC87744	ADC87744 DNA encod
12	1287.8	97.4	1901	12 ADO50297	ADO50297 Escherich
13	1287.8	97.4	1901	14 AED50819	AED50819 Escherich
14	1287.8	97.4	1901	15 AEE75418	AEE75418 E. coli B
15	1287	97.4	1299	3 ABK12514	ABK12514 DNA encod
16	1287	97.4	1299	8 ACA19297	ACA19297 Prokaryot
17	1287	97.2	1299	12 ADL16137	ADL16137 Escherich
18	1284.6	97.2	1901	12 ADO50295	ADO50295 Escherich

19	1284.6	97.2	1901	15 AEE75416	AEE75416 E. coli p
20	1283.8	97.1	1296	12 ADL16139	ADL16139 Shigella
21	1283	97.0	1901	12 ADO50296	ADO50296 Escherich
22	1283	97.0	1901	15 AEE75417	AEE75417 E. coli p
23	1282.2	97.0	3470	3 AAC68298	AAC68298 R15/APPA
24	1282.2	97.0	4060	3 AAC68296	AAC68296 R15/APPA
25	1282.2	97.0	6116	3 AAC68297	AAC68297 R15/APPA
26	1282.2	97.0	6708	3 AAC68295	AAC68295 R15/APPA
27	1282.2	97.0	17732	3 AAC68300	AAC68300 Lama2/App
28	1282.2	97.0	20623	3 AAC68294	AAC68294 Lama2/App
29	1281.4	96.9	1315	10 ADL02197	ADL02197 DNA encod
30	1280.6	96.9	1296	12 ADL16138	ADL16138 Shigella
31	1274.2	96.4	1399	15 AEF15218	AEF15218 Escherich
32	1270.2	96.1	5421	3 AAC68299	AAC68299 SH40/APPA
33	1267.8	95.9	1299	13 ADM76362	ADM76362 Phytase A
34	1265	95.7	1308	10 ADC87742	ADC87742 DNA encod
35	1265	95.7	1308	12 ADO50299	ADO50299 Escherich
36	1265	95.7	1308	15 AEE75420	AEE75420 E. coli B
37	1263.8	95.6	1489	3 AAA28216	AAA28216 E. coli a
38	1263.8	95.6	1489	8 ACC57672	ACC57672 Escherich
39	1263	95.5	1486	4 AAD06831	AAD06831 E. coli a
40	1262.2	95.5	1901	8 ADA19452	ADA19452 E. coli K
41	1257.4	95.1	1901	15 AEE75422	AEE75422 E. coli B
42	1255	94.9	1396	14 AED50817	AED50817 Escherich
43	1251.8	94.7	1486	4 AAD06832	AAD06832 E. coli a
44	1251.8	94.7	1486	8 ACC57673	ACC57673 Escherich
45	1217	92.1	1304	14 AED50824	AED50824 Escherich

ALIGNMENTS

RESULT 1	
ID AAC88885	standard; DNA; 1323 BP.
XX	
AC AAC88885;	
XX	
DT 07-MAR-2001 (first entry)	
XX	
DE Escherichia coli B phytase enzyme nucleotide sequence.	
XX	
KW Escherichia coli B; phytase enzyme; anabolic; phytate digestion;	
XX	
KM nutrition; ds.	
XX	
OS Escherichia coli.	
XX	
PN WO200071728-A1.	
XX	
PD 30-NOV-2000.	
XX	
PF 25-MAY-2000; 2000WO-US014846.	
XX	
PR 25-MAY-1999; 99US-00318528.	
XX	
PA (DIVE-) DIVERSA CORP.	
XX	
PI Short JM, Kretz KA;	
XX	
DR MPI: 2001-112081/12.	
XX	
PT P-PSDB; AAB37892.	
XX	
PT Improving the nutritional value of phytate-containing foodstuffs, using	
XX	
PT phytase enzymes which catalyze the liberation of inorganic phosphate from	
XX	
PS the phytates.	
XX	
PS Claim 2; Fig 1; 147pp; English.	
XX	
CC The present sequence encodes a phytase enzyme from Escherichia coli B.	
XX	
CC The enzyme catalyses the liberation of inorganic phosphate from the	
XX	
CC phytate in phytate-containing foodstuffs and can thus be used to improve	
XX	
CC the nutritional value of phytate rich ingredients	

SQ Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;
 Query Match 100.0%; Score 1322; DB 5; Length 1323;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 1 ATGAAAGGATCTTAAATCCATTTTATCTCTGATTCGGTTAAACCCGGAATCTGCA 60
Db 1 ATGAAAGGATCTTAAATCCATTTTATCTCTGATTCGGTTAAACCCGGAATCTGCA 60
Oy 61 TTCGTCAGAGTAGAGCCGGAAGCTGGAAGTGTGATTTGATTCAGTCGTATGAT 120
Db 61 TTCGTCAGAGTAGAGCCGGAAGCTGGAAGTGTGATTTGATTCAGTCGTATGAT 120
Oy 121 GTGCGTCTCCACCAAGGCCACGCACTGATGAGATGTCAACCCGAGAGCATGCGCA 180
Db 121 GTGCGTCTCCACCAAGGCCACGCACTGATGAGATGTCAACCCGAGAGCATGCGCA 180
Oy 181 ACCTGACCGGTAAACTGGGTTGCTGACACCGCGNGGTGGTGAAGTAACTATGCTATCTC 240
Db 181 ACCTGACCGGTAAACTGGGTTGCTGACACCGCGNGGTGGTGAAGTAACTATGCTATCTC 240
Oy 241 GGAACATTAACCAACGCGCAGCGCTCTGCTGAGCCGACGATTCCTGCGCAAAAGGCGTGC 300
Db 241 GGAACATTAACCAACGCGCAGCGCTCTGCTGAGCCGACGATTCCTGCGCAAAAGGCGTGC 300
Oy 301 CAGTCTGATCAGGTGCGATTAATTGCTGATGTCCAGACGCGTAAACCCGTTAAACAGGCCGA 360
Db 301 CAGTCTGATCAGGTGCGATTAATTGCTGATGTCCAGACGCGTAAACCCGTTAAACAGGCCGA 360
Oy 361 GCCTTCGCGCGCGGCTGGACCTGACTGTGCAATTAACCGTACATACCAGGAGATAGC 420
Db 361 GCCTTCGCGCGCGGCTGGACCTGACTGTGCAATTAACCGTACATACCAGGAGATAGC 420
Oy 421 TCCAGTCCCGATCCGTTAATTAACTCTTAATAAACTGCGGTTTGCCTGATTAACCGG 480
Db 421 TCCAGTCCCGATCCGTTAATTAACTCTTAATAAACTGCGGTTTGCCTGATTAACCGG 480
Oy 481 AACGTAATGACGCAATCTCTCAGCAGGCGGAGAGGTCAATTGCTGACTTTAACCGGCAAT 540
Db 481 AACGTAATGACGCAATCTCTCAGCAGGCGGAGAGGTCAATTGCTGACTTTAACCGGCAAT 540
Oy 541 CGGCAAAACGCGGTTTCGGAACCTGGAACGCGGCTTAAATTTTCCGCAATCAAACTTGTGC 600
Db 541 CGGCAAAACGCGGTTTCGGAACCTGGAACGCGGCTTAAATTTTCCGCAATCAAACTTGTGC 600
Oy 601 CTTAACCGTGAAGAAACAGGACGAAAGCTGTTCAATTACGACAGCATTAACATCGGAACCTC 660
Db 601 CTTAACCGTGAAGAAACAGGACGAAAGCTGTTCAATTACGACAGCATTAACATCGGAACCTC 660
Oy 661 AAGGTAAGCGCGCAATGCTCATTAACCGGTGCGTAAACGCTGCAATGCTGACG 720
Db 661 AAGGTAAGCGCGCAATGCTCATTAACCGGTGCGTAAACGCTGCAATGCTGACG 720
Oy 721 GAGATATTTCTCTGCAACAAGCAGAGGAAATGCGGAGCGGAGTGGGAAAGATCAAC 780
Db 721 GAGATATTTCTCTGCAACAAGCAGAGGAAATGCGGAGCGGAGTGGGAAAGATCAAC 780
Oy 781 GATTCAACACCAAGTGAACACTTGTGAAGTTTGTCAACGCGCAATTTTATTTGCTACAA 840
Db 781 GATTCAACACCAAGTGAACACTTGTGAAGTTTGTCAACGCGCAATTTTATTTGCTACAA 840
Oy 841 CGCAGCGCAAGAGTTGCGCGACGCGCGCAACCCGTTATTTGATTTGATTCGACG 900
Db 841 CGCAGCGCAAGAGTTGCGCGACGCGCGCAACCCGTTATTTGATTTGATTCGACG 900
Oy 901 TTGACGCCCCATCCACCGCAAAAACAGCGGTATGTGTGACATTAACCACTTCAGTACTG 960
Db 901 TTGACGCCCCATCCACCGCAAAAACAGCGGTATGTGTGACATTAACCACTTCAGTACTG 960
Oy 961 TTTATTTGCGGACACGATTAATCTGCGCAATCTCGGCGGCGCACTGGAAGCTCAACTGG 1020
Db 961 TTTATTTGCGGACACGATTAATCTGCGCAATCTCGGCGGCGCACTGGAAGCTCAACTGG 1020
  
```

```

Oy 1021 ACGCTTCCCGGTCAAGCCGGAATTAACAGCCCGCCAGGTGTGAACGTGTTGAACGCTGG 1080
Db 1021 ACGCTTCCCGGTCAAGCCGGAATTAACAGCCCGCCAGGTGTGAACGTGTTGAACGCTGG 1080
Oy 1081 CGTCGCTAAGCAGATTAACAGCAGTGAATTCAGGTTTCGCTGTCTTCAGACTTTACAG 1140
Db 1081 CGTCGCTAAGCAGATTAACAGCAGTGAATTCAGGTTTCGCTGTCTTCAGACTTTACAG 1140
Oy 1141 CAGATGCGGTATTAACGCGCTGTCAATTAAATACGCCGCCGAGAGGTGAACCTGACC 1200
Db 1141 CAGATGCGGTATTAACGCGCTGTCAATTAAATACGCCGCCGAGAGGTGAACCTGACC 1200
Oy 1201 CTGCGAGAGATGTGAAGACGAAATGCGAGGSCAATGTGTTGCTTGGACAGTTTAAACGCA 1260
Db 1201 CTGCGAGAGATGTGAAGACGAAATGCGAGGSCAATGTGTTGCTTGGACAGTTTAAACGCA 1260
Oy 1261 ATCGTGAATGAAGACGCAATCCGCGCTGACAGTTTGAGATCTCATCAACCATCAAC 1320
Db 1261 ATCGTGAATGAAGACGCAATCCGCGCTGACAGTTTGAGATCTCATCAACCATCAAC 1320
Oy 1321 TAA 1323
Db 1321 TAA 1323
  
```

RESULT 2
 AAD36473
 ID AAD36473 standard; DNA; 1323 BP.
 AC AAD36473;
 DT 21-AUG-2002 (first entry)
 XX Escherichia coli phytase DNA.
 KW Dietary aid; biocompatible composition; therapeutic; digestive tract;
 OS foodstuff; digestion; phytase; enzyme; gene; ds.
 XX Escherichia coli.
 FH Key Location/Qualifiers
 FT CDS 1..1323
 FT /tag=a
 FT /product="E. coli phytase protein"
 XX MO200189317-A2.
 XX 29-NOV-2001.
 XX 15-MAY-2001; 2001WO-US015764.
 XX 25-MAY-2000; 2000US-00580937.
 XX (DIVE-) DIVERSA CORP.
 XX Short JM, Kretz KA, O'donoghue E;
 XX WPI; 2002-164149/21.
 XX P-PSDB; AAE22836.
 PT New dietary aids comprising sustained release biocompatible compositions,
 PT comprise agent that assists in digestion, useful for delivering enzymes,
 PT therapeutics, medicine or agents to an organism.
 XX Disclousure; Fig 1; 89pp; English.
 CC The present invention relates to novel dietary aids comprising sustained
 CC release biocompatible composition which comprises an agent (enzymes such
 CC as phytase, amylase, esterase, protease) that assists in digestion. The
 CC biocompatible composition is effective upon oral consumption and release
 CC in the digestive tract of a subject. The dietary aids are useful for
 CC delivering enzymes, therapeutics, medicine and agents to an organism. The

CC use of enzymes and other agents in digestive aids of livestock or
 CC domesticated animals not only improves the animal's health and life
 CC expectancy but also assists in increasing the health of livestock or in
 CC the production of foodstuffs from livestock. The present sequence is
 CC Escherichia coli phytase DNA

XX Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;

Query Match 100.0%; Score 1322; DB 6; Length 1323;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGAAAGCATCTTAATCCATTTTATCTCTGATTCGGTTAACCCCGCAATCTGCA 60
DB 1 ATGAAAGCATCTTAATCCATTTTATCTCTCTGATTCGGTTAACCCCGCAATCTGCA 60
QY 61 TTGCGTCAAGTGAAGCCGAGCTGAAAGTGGAGTATGTCAGTGGTCAATGGT 120
DB 61 TTGCGTCAAGTGAAGCCGAGCTGAAAGTGGAGTATGTCAGTGGTCAATGGT 120
QY 121 GGGCGGCTCCCAACCAAGCCAGCACTGATGCAAGATGTCACCCCAAGCATGGCA 180
DB 121 GGGCGGCTCCCAACCAAGCCAGCACTGATGCAAGATGTCACCCCAAGCATGGCA 180
QY 122 GGGCGGCTCCCAACCAAGCCAGCACTGATGCAAGATGTCACCCCAAGCATGGCA 180
DB 122 GGGCGGCTCCCAACCAAGCCAGCACTGATGCAAGATGTCACCCCAAGCATGGCA 180
QY 181 ACCTGGCCGGTAAACTGGGTTGGCTGACACCGCGAGGTGGTAACTGGCTATCTC 240
DB 181 ACCTGGCCGGTAAACTGGGTTGGCTGACACCGCGAGGTGGTAACTGGCTATCTC 240
QY 241 GGACATTACCAACGCCAGCGTCTGGTAGCCGAGATTGTCGCAAAAAGGGCTGCGC 300
DB 241 GGACATTACCAACGCCAGCGTCTGGTAGCCGAGATTGTCGCAAAAAGGGCTGCGC 300
QY 301 CAGTCTGCTCAGAGTCCGATTTATGCTGATGTCAGACGCTAACCCGTTAAACAGCGCA 360
DB 301 CAGTCTGCTCAGAGTCCGATTTATGCTGATGTCAGACGCTAACCCGTTAAACAGCGCA 360
QY 361 GCGTTGCGCGCGGCTGACCTGACCTGACATACCGTACATACCGAGGAGATAGC 420
DB 361 GCGTTGCGCGCGGCTGACCTGACCTGACATACCGTACATACCGAGGAGATAGC 420
QY 421 TCCAGTCCCGATCCGTTATTTAATCCTTAATAAATCGGGCTTGCCAACTGGATTAACGCG 480
DB 421 TCCAGTCCCGATCCGTTATTTAATCCTTAATAAATCGGGCTTGCCAACTGGATTAACGCG 480
QY 481 AACGTAAGTGAAGGCACTCTGACAGAGGAGGAGGCTCAATTGCTAACTTACCGGCACT 540
DB 481 AACGTAAGTGAAGGCACTCTGACAGAGGAGGAGGCTCAATTGCTAACTTACCGGCACT 540
QY 541 CGGCAAAACGGCGTTTCGCGAATCGAAACGGGTGCTTAATTTCCGCAATCAAACTTGTGC 600
DB 541 CGGCAAAACGGCGTTTCGCGAATCGAAACGGGTGCTTAATTTCCGCAATCAAACTTGTGC 600
QY 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGACAGCATTAACATCGAACTC 660
DB 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGACAGCATTAACATCGAACTC 660
QY 661 AAGGTGAGGCGCGAATGCTCTTAATTAACCGGTGGGTAAAGCTGCAATGCTGACG 720
DB 661 AAGGTGAGGCGCGAATGCTCTTAATTAACCGGTGGGTAAAGCTGCAATGCTGACG 720
QY 721 GAGATATTTCTCTCTGCAACAGCAAGGAAATGCGGAGCCGGGGTGGGAAAGGATCAC 780
DB 721 GAGATATTTCTCTCTGCAACAGCAAGGAAATGCGGAGCCGGGGTGGGAAAGGATCAC 780
QY 781 GATTCAACACAGTGAACAATTGCTAAGTTTGCATACGCGCAATTTATTTGCTACAA 840
DB 781 GATTCAACACAGTGAACAATTGCTAAGTTTGCATACGCGCAATTTATTTGCTACAA 840
QY 841 CGCAGCGCCAGAGTTCCCGCGAGCGCGCAACCCGTTATTTGATTTGATTCAGTACGCG 900
DB 841 CGCAGCGCCAGAGTTCCCGCGAGCGCGCAACCCGTTATTTGATTTGATTCAGTACGCG 900
QY 901 TTGACGCCCATCAACCGCAAAAACAGGGGTATGGTGAATCAATTCACCTTCAGTACG 960
  
```

```

DB 901 TTGACGCCCATCAACCGCAAAAACAGGGGTATGGTGAATCAATTCACCTTCAGTACG 960
QY 961 TTTATTTGCGGACAGCATCTAATCTGCAAAATCTCGCGGCGCACTGAGCTCAACTG 1020
DB 961 TTTATTTGCGGACAGCATCTAATCTGCAAAATCTCGCGGCGCACTGAGCTCAACTG 1020
QY 1021 ACGCTTCCCGGTGAGCGGATTAACAGCGCGGAGGAGTGAATGCTGTTGAAGCTGG 1080
DB 1021 ACGCTTCCCGGTGAGCGGATTAACAGCGCGGAGGAGTGAATGCTGTTGAAGCTGG 1080
QY 1081 CGTGGCTTAAGGATTAACAGCAAGTGAATTCAGTTTCGCTGCTTCCAGACTTACAG 1140
DB 1081 CGTGGCTTAAGGATTAACAGCAAGTGAATTCAGTTTCGCTGCTTCCAGACTTACAG 1140
QY 1141 CAGATGCGTGAATTAACCGCGCTGCTAATTAATGCGCGCGGAGAGGTGAATGAC 1200
DB 1141 CAGATGCGTGAATTAACCGCGCTGCTAATTAATGCGCGCGGAGAGGTGAATGAC 1200
QY 1201 CTGGAGAGTGAAGAGGAAATGCGGAGGCAATGTTGCTGGCAGGTTTACGCA 1260
DB 1201 CTGGAGAGTGAAGAGGAAATGCGGAGGCAATGTTGCTGGCAGGTTTACGCA 1260
QY 1261 ATCGTAATGAAGACAGCATACCGCGGTCAGTTTGAAGTCTCATCACCATCAC 1320
DB 1261 ATCGTAATGAAGACAGCATACCGCGGTCAGTTTGAAGTCTCATCACCATCAC 1320
QY 1321 TAA 1323
DB 1321 TAA 1323
  
```

```

RESULT 3
AAD25460
ID AAD25460 standard; DNA; 1323 BP.
XX
AC AAD25460;
XX
DT 26-MAR-2002 (first entry)
XX
DE Escherichia coli B phytase DNA.
XX
KW Bacterial phytase; K12 appa phytase; proenzyme stability; anabolic;
KW gastrointestinal; nutritional value; feed treatment process; therapy;
KW thermal tolerance; growth performance; alcoholic drink; biopulping;
KW non-alcoholic drink; biobleaching; B phytase; ds.
XX
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT 1..1323
FT CDS /product= "E. coli B phytase protein"
XX
MO200190333-A2.
XX
PD 29-NOV-2001.
XX
PF 24-MAY-2001; 2001WO-US017118.
XX
PR 25-MAY-2000; 2000US-00580515.
XX
PA (DIVE-) DIVERSA CORP.
XX
PI Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, O'donoghue E;
XX MPI; 2002-083108/11.
XX DR P-PSDB; AAE15806.
XX
PT New bacterial phytase for e.g. improving the nutritional value of phytate
PT -containing foodstuffs and subsequently improving the growth performance
PT of an organism that consumes it, or in creating animal digestive systems.
XX
  
```


PR 01-MAR-1999; 99US-00259214.
 PR 13-APR-1999; 99US-00291931.
 PR 25-MAY-1999; 99US-00318528.
 PR 25-MAY-2000; 2000US-00580515.
 XX
 XX (SHOR/) SHORT J M.
 PA (KRET/) KRETZ K A.
 PA (GRAY/) GRAY K A.
 PA (BART/) BARTON N R.
 PA (GARR/) GARRETT J B.
 PA (DONO/) O' DONOGHUE E.
 PA (MATH/) MATHUR E J.
 XX
 PI Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB;
 PI O' Donoghue E, Mathur EJ;
 XX
 DR WPI; 2003-040002/03.
 XX P-PSDB; ADA19446.
 XX
 PT Isolated *Escherichia coli* polynucleotide encoding a modified phytase
 PT enzyme, useful in the production of animal feed, for improving the
 PT nutritional value of phytate-containing feedstuff and for enhancing
 PT digestion in humans and animals.
 XX
 XX
 PS Claim 2; Fig 1; 62pp; English.

XX The invention relates to an isolated *Escherichia coli* polynucleotide
 CC encoding a phytase enzyme appearing as ADA19450 and having amino acids
 CC modification from W68E, Q84W, A95P, K97C, S168E, R180Y, N226C or Y277D.
 CC Also included from the *E. coli* appa gene ADA19449 (or an oligonucleotide
 CC derived from it) or its mutant sequence ADA19452, expression vectors,
 CC host cells, a method of improving nutritional value of a phytate-
 CC containing feedstuff by contacting the phytate-containing feedstuff with
 CC a pure phytase enzyme mutated phytase (where the phytase enzyme catalyzes
 CC the liberation of inorganic phosphate from the phytate in the phytate-
 CC containing feedstuff), a method to produce an animal feed containing a
 CC microbial phytase (comprising culturing the plant cell, plant part or
 CC plant under conditions where the nucleotide sequence is expressed and
 CC converting the plant cells, plants or plant into a composition for animal
 CC feed), a feed composition for animals (comprising the plant seeds, plant
 CC cells, plant parts or plants in admixture with a phytate-containing
 CC feedstuff), a method to treat a human or an animal able to benefit from
 CC digestive enhancement by the activity of an exogenous phytase enzymes
 CC comprising administering to the human or animal the plant seed, plant
 CC cells, plant parts or plants of a transgenic plant which is modified to
 CC contain an expression system which expresses a nucleotide sequence
 CC encoding a phytase enzyme, a transgenic non-human organism whose genome
 CC comprising a heterologous nucleic acid sequence encoding a polypeptide
 CC having phytase activity. The phytase enzyme is useful for improving the
 CC nutritional value of phytate-containing feedstuff, in the production of
 CC animal feed and for enhancing digestion in humans and animals. The
 CC invented method improves thermal tolerance and protease stability. It
 CC also improves the feeding value of phytate rich ingredients. The present
 CC sequence encodes *E. coli* B wild-type phytase.
 XX

XX Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;

Query Match 100.0%; Score 1322; DB 8; Length 1323;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAGCATCTTAATCCCATTTTATCTCTTGATTCGGTTAACCCCGCAATTCGA 60
 DB 1 ATGAAAGCATCTTAATCCCATTTTATCTCTTGATTCGGTTAACCCCGCAATTCGA 60
 QY TTTCGCTCAGATGAGCGCGAGCTGAAGCTGAAGTGTGTGATTTGTCACTGTCATGGT 120
 DB 61 TTTCGCTCAGATGAGCGCGAGCTGAAGTGTGTGATTTGTCACTGTCATGGT 120
 QY TTTCGCTCAGATGAGCGCGAGCTGAAGTGTGTGATTTGTCACTGTCATGGT 120
 DB 61 TTTCGCTCAGATGAGCGCGAGCTGAAGTGTGTGATTTGTCACTGTCATGGT 120
 QY GTGCGGTCTTCAACCAAGGCCAAGCACTGATGCAAGATGTCACCCCAAGCATGGCCA 180
 DB 121 GTGCGGTCTTCAACCAAGGCCAAGCACTGATGCAAGATGTCACCCCAAGCATGGCCA 180
 QY GTGCGGTCTTCAACCAAGGCCAAGCACTGATGCAAGATGTCACCCCAAGCATGGCCA 180
 DB 121 GTGCGGTCTTCAACCAAGGCCAAGCACTGATGCAAGATGTCACCCCAAGCATGGCCA 180

QY 181 ACTGGCGGGTAAACTGGGTTGGCTGACACCGGNGGGTGAAGTAACTGGCTATCTC 240
 DB 181 ACTGGCGGGTAAACTGGGTTGGCTGACACCGGNGGGTGAAGTAACTGGCTATCTC 240
 QY 241 GGACATTAACAACGCGCAGCGCTGTGTAGCGCAGATTGCTGGCGAAAAAGGGCTGCCG 300
 DB 241 GGACATTAACAACGCGCAGCGCTGTGTAGCGCAGATTGCTGGCGAAAAAGGGCTGCCG 300
 QY 301 CAGTGTGTGAGGTGTGCTGATTAATGCTGATGTCAGACGCTAACCCGTAAACAGCGCAA 360
 DB 301 CAGTGTGTGAGGTGTGCTGATTAATGCTGATGTCAGACGCTAACCCGTAAACAGCGCAA 360
 QY 361 GCGTTCGCGCGGGGTGGCCTGACCTGATGTAATACCTGATCAATCCAGGAGATAG 420
 DB 361 GCGTTCGCGCGGGGTGGCCTGACCTGATGTAATACCTGATCAATCCAGGAGATAG 420
 QY 421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAACATGCGCTTGGCACTGTAAACCGG 480
 DB 421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAACATGCGCTTGGCACTGTAAACCGG 480
 QY 481 AAGTGAATGACGCGATCTCTAGCAGGGCAGAGGGTCAATTTGCTGACTTTAACGGGCAT 540
 DB 481 AAGTGAATGACGCGATCTCTAGCAGGGCAGAGGGTCAATTTGCTGACTTTAACGGGCAT 540
 QY 541 CGGCAAAACGCGCTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAACTTGTCC 600
 DB 541 CGGCAAAACGCGCTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAACTTGTCC 600
 QY 601 CTTAAACGTGAAGAAACAGGACGAAAGCTGTCTAATTAACGACGACATTAACATCGAATC 660
 DB 601 CTTAAACGTGAAGAAACAGGACGAAAGCTGTCTAATTAACGACGACATTAACATCGAATC 660
 QY 661 AAGTGAAGCGCGACATGCTCTAATTAACCGGTGCGTAAAGCTTCGATCAATGCTGACG 720
 DB 661 AAGTGAAGCGCGACATGCTCTAATTAACCGGTGCGTAAAGCTTCGATCAATGCTGACG 720
 QY 721 GAGATATTTCTCTGCAACAGCAACAGGAAATGCCGAGCCGGGTGGGAAAGATCACC 780
 DB 721 GAGATATTTCTCTGCAACAGCAACAGGAAATGCCGAGCCGGGTGGGAAAGATCACC 780
 QY 781 GATTCAACACAGTGAACACCTTGTGATTTGATTAACGCGCAATTTTATTTGCTACAA 840
 DB 781 GATTCAACACAGTGAACACCTTGTGATTTGATTAACGCGCAATTTTATTTGCTACAA 840
 QY 841 CGCAGCGCAGAGTGTCCCGCAGCGCGCACCCCGTATTGATTTGATTCAGCAGCG 900
 DB 841 CGCAGCGCAGAGTGTCCCGCAGCGCGCACCCCGTATTGATTTGATTCAGCAGCG 900
 QY 901 TTGAGCGCCCATCCACCGCAAAAACAGGCGTATGCTGACATTACCACTTCAGTACG 960
 DB 901 TTGAGCGCCCATCCACCGCAAAAACAGGCGTATGCTGACATTACCACTTCAGTACG 960
 QY 961 TTTATTTGCGGACACAGTACTTAATCTGGCAATCTCGCGGCGCATCTGAGCTCAACTGG 1020
 DB 961 TTTATTTGCGGACACAGTACTTAATCTGGCAATCTCGCGGCGCATCTGAGCTCAACTGG 1020
 QY 1021 ACGCTTCCCGGTACGCGCGGATTAACAGCGCGCGAGTGTGTAATCTGTGTTTGAACGCTGG 1080
 DB 1021 ACGCTTCCCGGTACGCGCGGATTAACAGCGCGCGAGTGTGTAATCTGTGTTTGAACGCTGG 1080
 QY 1081 CGTGGCTTAAGGATTAACAGCAGTGTGATTCAGGTTTGGCTGTCTTCCAGACTTTACG 1140
 DB 1081 CGTGGCTTAAGGATTAACAGCAGTGTGATTCAGGTTTGGCTGTCTTCCAGACTTTACG 1140
 QY 1141 CAGATGCGGTATTAACAGCGCGCTGTCTAATTAATACGCGCGCGAGAGGTGAACTGAC 1200
 DB 1141 CAGATGCGGTATTAACAGCGCGCTGTCTAATTAATACGCGCGCGAGAGGTGAACTGAC 1200
 QY 1201 CTGGCAGATGTGAAGACGAAATGCGCAGGCGCATGTGTTGTTGCAAGTTTTACGCA 1260
 DB 1201 CTGGCAGATGTGAAGACGAAATGCGCAGGCGCATGTGTTGTTGCAAGTTTTACGCA 1260
 QY 1261 ATCGTGAATGAAGCAGCATACCGGCGTGCAGTTTGAATCTCATCAGATCAGATCAC 1320

Db 1261 ATCCGATGAAGACGCAATCCGGCGGCAATTGATCTCATCAATCACCATCACC 1320
Qy 1321 TAA 1323
Db 1321 TAA 1323

RESULT 5
AD050291
ID AD050291 standard; DNA; 1323 BP.
AC AD050291;
XX 29-JUN-2004 (first entry)
DE Escherichia coli B phytase DNA.
XX Phytase; bacteria; thermal tolerance; protease stability; foodstuff;
KW phytase; animal feed; fish feed; dough; baking; gene; ds.
XX Escherichia coli; B.
OS
XX Key Location/Qualifiers
FH 1..1323
FT CDS /tag= a
FT /product= "Phytase enzyme"
FT 214..216
FT /tag= b
FT /note= "Encodes Arg"

US2004091968-A1.
XX 13-MAY-2004.
XX 20-JUN-2003; 2003US-00601319.
XX 13-AUG-1997; 97US-00910798.
XX 01-MAR-1999; 99US-00259214.
XX 13-APR-1999; 99US-00291931.
XX 25-MAY-1999; 99US-00318528.
XX 25-MAY-2000; 2000US-00580515.
XX 24-MAY-2001; 2001US-00866379.
XX (SHORT) SHORT J M.
XX (KRETT) KRETT K.
XX (GRAY) GRAY K A.
XX (BART) BARTON N R.
XX (GAR) GARRETT J B.
XX (ODON) O'DONOGHUE B.
XX (MATH) MATHER E J.
XX Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E,
XX Mather EJ;
XX MPI: 2004-374952/35.
XX P-PSDB; AD050292.
XX Producing phytase, involves providing nucleic acid derived from bacteria
XX encoding polypeptide having phytase activity, and expressing nucleic acid
XX in yeast.
XX Claim 3; SEQ ID NO 1; 74pp; English.

XX The invention relates to a method for producing phytase that involves
XX providing a nucleic acid encoding phytase derived from a bacteria, and
XX expressing the nucleic acid in a yeast under conditions that allow
XX expression of the enzyme in the yeast. The invention also relates to
XX modified phytase enzyme which has improved thermal tolerance and protease
XX stability at low pH. The phytase enzyme can be used in foodstuffs to
XX improve the feeding value of phytate rich ingredients, and in diet of
XX numerous animals including mammals, fowls and fishes, commercially
XX significant mammals such as pigs, goats, laboratory rodents, commercially

CC significant avian species such as chicken, ducks, doves, parrot, etc.,
CC commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC etc., in dough making and baking, in dietary aids for animals. The method
CC provides easy manufacture of the active ingredient loaded biocompatible
CC composition, higher yields and loading efficiency. The phytase
CC incorporated in to the dietary aid is safe for animals. The present
CC sequence is Escherichia coli B phytase DNA.
XX
SQ Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;
Query Match 100.0%; Score 1322; DB 12; Length 1323;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAAGGAGATCTTAATCCATTTTATCTTGTGATTCGTTAACCCGCAATTCGCA 60
Db 1 ATGAAGGAGATCTTAATCCATTTTATCTTGTGATTCGTTAACCCGCAATTCGCA 60
Qy TTCCGTCAGAGTGAAGCCGAGCTGAGCTGGAAGTGTGATTCAGTCATGCT 120
Db TTCCGTCAGAGTGAAGCCGAGCTGAGCTGGAAGTGTGATTCAGTCATGCT 120
Qy 121 GTGCGTGTCTCAACCAAGCCCAAGCACTGATGCAAGATGTACCCAGACATGGCCA 180
Db 121 GTGCGTGTCTCAACCAAGCCCAAGCACTGATGCAAGATGTACCCAGACATGGCCA 180
Qy 181 ACCTGGCCGGTAAACTGGGTGGCTGACACCGGAGGTGGTGAATTCGCTATCTC 240
Db 181 ACCTGGCCGGTAAACTGGGTGGCTGACACCGGAGGTGGTGAATTCGCTATCTC 240
Qy 241 GGACATTACCAACCCGACGCTGTGTGAGCCGACGATTTGTGCGGAAAAAGGCTGCCG 300
Db 241 GGACATTACCAACCCGACGCTGTGTGAGCCGACGATTTGTGCGGAAAAAGGCTGCCG 300
Qy 301 CAGTCTGTGACGTGCGGATTTGTGATGTGACAGAGCTACCCGTTAAACGCGGAA 360
Db 301 CAGTCTGTGACGTGCGGATTTGTGATGTGACAGAGCTACCCGTTAAACGCGGAA 360
Qy 361 GCGTTCGCGCGCGGCGTGGCACTGACGTGCAATTAACCGTACCCGAGGAGATACG 420
Db 361 GCGTTCGCGCGCGGCGTGGCACTGACGTGCAATTAACCGTACCCGAGGAGATACG 420
Qy 421 TCCAGTCCCGATCCGTTATTTAATCTCTAATAAACTGCGCTTTCGCAATGATACGCG 480
Db 421 TCCAGTCCCGATCCGTTATTTAATCTCTAATAAACTGCGCTTTCGCAATGATACGCG 480
Qy 481 AACGTGACTGACGCGATCTCAGCAGGAGGCTCAATTCGCTTAACCGGCGAT 540
Db 481 AACGTGACTGACGCGATCTCAGCAGGAGGCTCAATTCGCTTAACCGGCGAT 540
Qy 541 CGGAAACGGCGTTCGCGAATCGGAAACGGGTGCTTAATTTTCGCGAATCAACTTGTGC 600
Db 541 CGGAAACGGCGTTCGCGAATCGGAAACGGGTGCTTAATTTTCGCGAATCAACTTGTGC 600
Qy 601 CTTAAACGTGAGAAACAGAGCAAGAGCTGTTCAATTAACGACGATTAACATCCGAACTC 660
Db 601 CTTAAACGTGAGAAACAGAGCAAGAGCTGTTCAATTAACGACGATTAACATCCGAACTC 660
Qy 661 AAGGTGAGCGCGCAATGTCTCAATTAACCGGTGCGGTAACCTCGCATCATCTGACG 720
Db 661 AAGGTGAGCGCGCAATGTCTCAATTAACCGGTGCGGTAACCTCGCATCATCTGACG 720
Qy 721 GAGATATTTCTCTGCAACCAAGCAAGGAAATGCGGAGCGGGGTGGGAAAGATCAC 780
Db 721 GAGATATTTCTCTGCAACCAAGCAAGGAAATGCGGAGCGGGGTGGGAAAGATCAC 780
Qy 781 GATTCACCAACGATGGAACACCTTCTAATGTTGATTAACGCGCATTTTATTTCTACAA 840
Db 781 GATTCACCAACGATGGAACACCTTCTAATGTTGATTAACGCGCATTTTATTTCTACAA 840
Qy 841 CGCAGCGCAGAGTGTGCGCGGACGCGCACCCCGTTATTGATTTGATGATGACGCG 900
Db 841 CGCAGCGCAGAGTGTGCGCGGACGCGCACCCCGTTATTGATTTGATGATGACGCG 900

```

QY 901 TTGAGCGCCCATCCACGCGAATAAGGCGTATGTTGTCATTACCACTTCAGTACG 960
Db 901 TTGAGCGCCCATCCACGCGAATAAGGCGTATGTTGTCATTACCACTTCAGTACG 960
QY 961 TTTATTTGCGGACACGATATCTAATCTGGCAATCTCGGCGGCGACATGAGCTCAACTG 1020
Db 961 TTTATTTGCGGACACGATATCTAATCTGGCAATCTCGGCGGCGACATGAGCTCAACTG 1020
QY 1021 ACCGTTCCCGGTGACCGCGATTAACAGCCGCCGACGTTGTAATCTGTTTGAACCTG 1080
Db 1021 ACCGTTCCCGGTGACCGCGATTAACAGCCGCCGACGTTGTAATCTGTTTGAACCTG 1080
QY 1081 CGTGGCTTAAGCATTAACAGCCGATGATTCAGTTTCTGTTCTTCCAGACTTTACG 1140
Db 1081 CGTGGCTTAAGCATTAACAGCCGATGATTCAGTTTCTGTTCTTCCAGACTTTACG 1140
QY 1141 CAGATCGTATTAATAAGCGCGCTGTCTAATTAATACCGCCGCGAGAGGTGAATGAC 1200
Db 1141 CAGATCGTATTAATAAGCGCGCTGTCTAATTAATACCGCCGCGAGAGGTGAATGAC 1200
QY 1201 CTGGCAGATGTAAGAGCGAAATGCGCAGGCGATGTTGTTGCGAGGTTTACGCA 1260
Db 1201 CTGGCAGATGTAAGAGCGAAATGCGCAGGCGATGTTGTTGCGAGGTTTACGCA 1260
QY 1261 ATGTGAATGAAGACGAGCATACCGGCGTGCAGTTTGATTCATACCATCACCTCAC 1320
Db 1261 ATGTGAATGAAGACGAGCATACCGGCGTGCAGTTTGATTCATACCATCACCTCAC 1320
QY 1321 TAA 1323
Db 1321 TAA 1323

```

RESULT 6
ID ABE75412
ABE75412 standard; DNA; 1323 BP.

AC ABE75412;
XX
DT 23-FEB-2006 (first entry)
XX
DE E. coli B phytase coding sequence SEQ ID NO: 1.
XX
KW ds; coding sequence; thermostable; phytase; protein stabilization;
KW pharmaceutical; fertilizer; cramp; cell culture; osteoporosis;
XX osteopathic; food.
OS Escherichia coli B.
XX
XX Key Location/Qualifiers
FH 1.1323
FT /tag= a
FT /product= "Phytase"
FT 214..216
FT unsure
FT /tag= b
FT /note= "ecodes Arg"

US2005281792-A1.
PN
XX
PD 22-DEC-2005.
XX
PF 01-SEP-2004; 2004US-00931115.
XX
PR 13-AUG-1997; 97US-00910798.
PR 01-MAR-1999; 99US-00259214.
PR 13-APR-1999; 99US-00291931.
PR 25-MAY-1999; 99US-00318528.
PR 25-MAY-2000; 2000US-00580515.
PR 24-MAY-2001; 2001US-00866379.
XX
XX (SHOR/) SHORT J M.
PA (KRETT) KRETZ K A.

```

PA (GRAY/) GRAY K A.
PA (BART/) BARTON N R.
PA (GARR/) GARRETT J B.
PA (ODON/) O'DONOGHUE E.
PA (BAUM/) BAUM W.
PA (ROBE/) ROBERTSON D E.
PA (ZORN/) ZORNER P.
PI Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, Odonoghue E,
PI Baum W, Robertson DE, Zorner P;
XX WPI; 2006-055022/06.
XX P-PSDB; ABE75413.
XX
XX DR
XX PT Formulation useful as dietary supplement for treating, preventing or
XX PT reversing osteoporosis and bone loss, and preventing muscle cramps,
XX PT comprises one or more polypeptides having phytase activity.
XX
XX PS Claim 1; SEQ ID NO 1; 82bp; English.
XX
XX CC The present sequence is that of a polynucleotide encoding the Escherichia
XX CC coli B phytase. The present invention relates to a novel formulation,
XX CC comprising at least one polypeptide having phytase activity, where the
XX CC polypeptide is a modified sequence (ABE75421) derived from the
XX CC Escherichia coli K-12 appa phytase (ABE75419). The modification of the
XX CC enzyme, comprising a series of mutations, improves the thermal tolerance
XX CC and protease stability of the protein. The specification also claims a
XX CC pharmaceutical composition; a kit; an immobilized phytase; a fertilizer
XX CC or soil additive; a liquid supplement for preventing muscle cramps; a
XX CC hydrating agent; a tissue culture or cell culture media; and a plant food
XX CC additive, all comprising the novel phytase, and a method of reducing
XX CC pollution and increasing nutrient availability in an environment or
XX CC environmental sample by degrading environmental phytic acid. The novel
XX CC enzyme is a dietary supplement useful for treating, preventing or
XX CC reversing osteoporosis or bone loss, and preventing muscle cramps. The
XX CC liquid supplement is useful for preventing muscle cramps. The method is
XX CC useful for reducing pollution and increasing nutrient availability in an
XX CC environment or environmental sample by degrading environmental phytic
XX CC acid, where the environment or environmental sample comprises a soil or a
XX CC body of water. The immobilized phytase is useful in foodstuffs for
XX CC improving the feeding value of phytate rich ingredients.
XX
XX SQ Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;

```

Query Match 100.0%; Score 1322; DB 15; Length 1323;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGAAGGATCTTAATCCATTTTATCTCTTGATTCCTTAACCCCGCAATCTGA 60
Db 1 ATGAAGGATCTTAATCCATTTTATCTCTTGATTCCTTAACCCCGCAATCTGA 60
QY 61 TTGCGTCAGAGTACCGGAGCTGAAGCTGGAATGTTGATGTCAGTGCATGAT 120
Db 61 TTGCGTCAGAGTACCGGAGCTGAAGCTGGAATGTTGATGTCAGTGCATGAT 120
QY 121 GTGCGTGTCTCAACCAAGCGCACGCACTGATGACAGTGTCAACCCAGACGATGCGCA 180
Db 121 GTGCGTGTCTCAACCAAGCGCACGCACTGATGACAGTGTCAACCCAGACGATGCGCA 180
QY 181 ACCTGGCCGGTAAACTGGGTTGGCTGACACCGCGAGGTGTGACTAATGCGCTATCTC 240
Db 181 ACCTGGCCGGTAAACTGGGTTGGCTGACACCGCGAGGTGTGACTAATGCGCTATCTC 240
QY 241 GGAACATTACCAAGCGCAGCTGTGGTAGCCGACGATGTCGCAAAAAGGCGTCCG 300
Db 241 GGAACATTACCAAGCGCAGCTGTGGTAGCCGACGATGTCGCAAAAAGGCGTCCG 300
QY 301 CAGTCTGATCAGGTGCGCATTAATGCTGATGTCAGACGCGTACCCGTAACAGCGCAA 360
Db 301 CAGTCTGATCAGGTGCGCATTAATGCTGATGTCAGACGCGTACCCGTAACAGCGCAA 360
QY 361 GCGTTCGCGCGCGGCTGACCTGATCTGTGAATTAACCGTAACATCCAGGCAATACG 420

```

```

Db      ||||| 361 GCCTTCGCGCGCGGCTGGCACTGACTGTGCATTAACCGTACATACCAAGGCAATACG 420
Qy      ||||| 421 TCACGTCCCGATCCGTTATTTATTCCTCTAAAACCTGGCTTGGCAACGTGANACGCG 480
Db      ||||| 421 TCACGTCCCGATCCGTTATTTATTCCTCTAAAACCTGGCTTGGCAACGTGANACGCG 480
Qy      ||||| 481 AACGTGACGTGACGATCCTGACGAGGGGAGAGGGTCAATTGCTGACCTTTACCGGGCAT 540
Db      ||||| 481 AACGTGACGTGACGATCCTGACGAGGGGAGAGGGTCAATTGCTGACCTTTACCGGGCAT 540
Qy      ||||| 541 CGGCAACGCGCTTTCGCGAACGTGACGCGGTCTTAAATTTTCGCGCAATCAAACTTGTGC 600
Db      ||||| 541 CGGCAACGCGCTTTCGCGAACGTGACGCGGTCTTAAATTTTCGCGCAATCAAACTTGTGC 600
Qy      ||||| 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGACGACATTAACATCGGAATC 660
Db      ||||| 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGACGACATTAACATCGGAATC 660
Qy      ||||| 661 AAGGTAGGCGCGACATGCTCATTAACCGGTGGGTGAAGCTGCAATGCTGACG 720
Db      ||||| 661 AAGGTAGGCGCGACATGCTCATTAACCGGTGGGTGAAGCTGCAATGCTGACG 720
Qy      ||||| 721 GAGATATTTCTCTGCAACAAAGCAGGAGATGCGGAGCGCGGGTGGGAAAGATCAC 780
Db      ||||| 721 GAGATATTTCTCTGCAACAAAGCAGGAGATGCGGAGCGCGGGTGGGAAAGATCAC 780
Qy      ||||| 781 GATTACACACGATGGAACACTTGTGTAAGTTGCAATACGCGCAATTTATTTGCTACA 840
Db      ||||| 781 GATTACACACGATGGAACACTTGTGTAAGTTGCAATACGCGCAATTTATTTGCTACA 840
Qy      ||||| 841 CGCAGCGCAGAGTTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
Db      ||||| 841 CGCAGCGCAGAGTTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
Qy      ||||| 901 TTGAGCGCCCATCAACCGCAAAAAGAGCGTATGTTGACCTTACCCACTTCACTACTG 960
Db      ||||| 901 TTGAGCGCCCATCAACCGCAAAAAGAGCGTATGTTGACCTTACCCACTTCACTACTG 960
Qy      ||||| 961 TTTATTTGCGGACAGATATCTATCTGCGCAATCTCGGCGGCGGCGGCGGCGGCGG 1020
Db      ||||| 961 TTTATTTGCGGACAGATATCTATCTGCGCAATCTCGGCGGCGGCGGCGGCGGCGG 1020
Qy      ||||| 1021 ACGCTTCCCGTCAACCGGATTAACGCGCGGCGGAGGTGTAAGTGTGTAAGCGTGG 1080
Db      ||||| 1021 ACGCTTCCCGTCAACCGGATTAACGCGCGGCGGAGGTGTAAGTGTGTAAGCGTGG 1080
Qy      ||||| 1081 CGTCCGCTAAGCGATTAACGCGGATGATTCAGTTTCGCTGCTTCCAGACTTTAAG 1140
Db      ||||| 1081 CGTCCGCTAAGCGATTAACGCGGATGATTCAGTTTCGCTGCTTCCAGACTTTAAG 1140
Qy      ||||| 1141 CAGATCCGTGATTAACCGCGCTGTGATTAATACGCGCGCGGAGAGGTGAACCTGACC 1200
Db      ||||| 1141 CAGATCCGTGATTAACCGCGCTGTGATTAATACGCGCGCGGAGAGGTGAACCTGACC 1200
Qy      ||||| 1201 CTGGAGAGATGGAAGAGGAATGCGGAGGCGATGTTGCTGTGGCAGGTTTACGCA 1260
Db      ||||| 1201 CTGGAGAGATGGAAGAGGAATGCGGAGGCGATGTTGCTGTGGCAGGTTTACGCA 1260
Qy      ||||| 1261 ATTCGTAATGAAGACGCGATACCGGCGGTGCAATTCATACCATCACCATGAC 1320
Db      ||||| 1261 ATTCGTAATGAAGACGCGATACCGGCGGTGCAATTCATACCATCACCATGAC 1320
Qy      ||||| 1321 TAA 1323
Db      ||||| 1321 TAA 1323

```

```

XX      ||||| 29-JUN-2004 (first entry)
DT      |||||
XX      ||||| Kangaroo rat Escherichia coli phytase 872PHI mutant DNA.
DE      |||||
XX      ||||| Phytase; bacteria; thermal tolerance; protease stability; foodstuff;
KW      ||||| phytase; animal feed; fish feed; dough; baking; mutant; gene; ds.
XX      ||||| Escherichia coli.
OS      |||||
XX      |||||
FH      ||||| Key
FT      ||||| Location/Qualifiers
FT      ||||| 1..1323
FT      ||||| /*tag= a
FT      ||||| /product= "Phytase mutant enzyme"
FT      ||||| 214..216
FT      ||||| /*tag= b
FT      ||||| /note= "Encodes Arg"
FT      ||||| replace(527,A)
FT      ||||| /*tag= c
FT      ||||| replace(893,T)
FT      ||||| /*tag= d
FT      ||||| mutation
FT      ||||| replace(895,G)
FT      ||||| /*tag= e
FT      ||||| mutation
FT      ||||| replace(934,G)
FT      ||||| /*tag= f
XX      |||||
PN      ||||| US2004091968-A1.
XX      |||||
PD      ||||| 13-MAY-2004.
XX      |||||
XX      ||||| 20-JUN-2003; 2003US-00601319.
XX      |||||
XX      ||||| 13-AUG-1997; 97US-00910798.
PR      ||||| 01-MAR-1999; 99US-00259214.
PR      ||||| 13-APR-1999; 99US-00291931.
PR      ||||| 25-MAY-1999; 99US-00318528.
PR      ||||| 25-MAY-2000; 2000US-00580515.
PR      ||||| 24-MAY-2001; 2001US-00866379.
XX      |||||
PA      ||||| (SHORT) SHORT J M.
PA      ||||| (KRETZ) KRETZ K.
PA      ||||| (GRAY) GRAY K A.
PA      ||||| (BART) BARTON N R.
PA      ||||| (GARR) GARBETT J B.
PA      ||||| (ODON) O'DONOGHUE E.
PA      ||||| (MATH) MATHER E J.
XX      |||||
XX      ||||| Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI      ||||| Mather EJ;
PI      |||||
DR      ||||| WPI, 2004-374952/35.
DR      ||||| P-PSDB; ADO50302.
XX      |||||
XX      ||||| Producing phytase, involves providing nucleic acid derived from bacteria
PT      ||||| encoding polypeptide having phytase activity, and expressing nucleic acid
PT      ||||| in yeast.
XX      |||||
PS      ||||| Disclosure; Page: 74pp; English.
XX      |||||
XX      ||||| The invention relates to a method for producing phytase that involves
CC      ||||| providing a nucleic acid encoding phytase derived from a bacteria, and
CC      ||||| expressing the nucleic acid in a yeast under conditions that allow
CC      ||||| expression of the enzyme in the yeast. The invention also relates to
CC      ||||| modified phytase enzyme which has improved thermal tolerance and protease
CC      ||||| stability at low pH. The phytase enzyme can be used in foodstuffs to
CC      ||||| improve the feeding value of phytase rich ingredients, and in diet of
CC      ||||| numerous animals including mammals, fowls and fishes, commercially
CC      ||||| significant avian species such as pigs, goats, laboratory rodents, commercially
CC      ||||| significant farmed fish such as guppy, zebrafish, mollly, swordtail,
CC      ||||| etc., in dough making and baking, in dietary aids for animals. The method
CC      ||||| provides easy manufacture of the active ingredient loaded biocompatible
CC      ||||| composition, higher yields and loading efficiency. The phytase

```

CC Incorporated in to the dietary aid is safe for animals. The present
 CC sequence is Kangaroo rat *Escherichia coli* phytase mutant DNA. Note: The
 CC present sequence is not shown in the specification but has been derived
 CC from *Escherichia coli* B phytase DNA AD050291.

XX Sequence 1323 BP, 325 A, 353 C, 356 G, 288 T, 0 U, 1 Other;

Query Match 99.5%; Score 1315.6; DB 12; Length 1323;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1319; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 ATGAAAGCGATCTTAATCCATTTTATCTCTTCTGATTCGTTAACCCGCAATCTGCA 60
DB 1 ATGAAAGCGATCTTAATCCATTTTATCTCTTCTGATTCGTTAACCCGCAATCTGCA 60
QY 61 TTCGCTCAGATGAGCGCGAGCTGGAAGCTGGAAGAGTGTGTATGTGCAATGCTCATGTG 120
DB 61 TTCGCTCAGATGAGCGCGAGCTGGAAGCTGGAAGAGTGTGTATGTGCAATGCTCATGTG 120
QY 121 GTGCGTCTCAACCAAGGCCCAAGCTGATGAGATGTCAACCCAGACGATGGCCA 180
DB 121 GTGCGTCTCAACCAAGGCCCAAGCTGATGAGATGTCAACCCAGACGATGGCCA 180
QY 181 ACCTGGCCGGTAAACTGGGTTGGCTGACACCGCGAGTGTGTAGCTAATCGCTATCTC 240
DB 181 ACCTGGCCGGTAAACTGGGTTGGCTGACACCGCGAGTGTGTAGCTAATCGCTATCTC 240
QY 241 GGAACATTAACCAAGCCAGCTCTGTGTAGCCGACGATGCTGGCGAAAAAGGCTCCCG 300
DB 241 GGAACATTAACCAAGCCAGCTCTGTGTAGCCGACGATGCTGGCGAAAAAGGCTCCCG 300
QY 301 CAGTCTGTGAGGTCCGGAATTATGCTGATGTGAGAGGATACCCGTAACAGAGCGAA 360
DB 301 CAGTCTGTGAGGTCCGGAATTATGCTGATGTGAGAGGATACCCGTAACAGAGCGAA 360
QY 361 GCCTTGGCCGCGCGGCTGGCACTGTGCAATAACCTACATACCCAGCGAGATAG 420
DB 361 GCCTTGGCCGCGCGGCTGGCACTGTGCAATAACCTACATACCCAGCGAGATAG 420
QY 421 TCCAGTCCCGATTCGTTATTTAATCTCTTAATACTGGGTTGCCAATGATACGG 480
DB 421 TCCAGTCCCGATTCGTTATTTAATCTCTTAATACTGGGTTGCCAATGATACGG 480
QY 481 AACGTGACTGACGCGATCTCTGACAGAGGAGGATCAATTGCTTAACCGGCGAT 540
DB 481 AACGTGACTGACGCGATCTCTGACAGAGGAGGATCAATTGCTTAACCGGCGAT 540
QY 541 CGGCAAAACGCGCTTTCGCAACTGGAACGCGTCTTAATTTTCGCAATCAAACTGTGC 600
DB 541 CGGCAAAACGCGCTTTCGCAACTGGAACGCGTCTTAATTTTCGCAATCAAACTGTGC 600
QY 601 CTTTAAACGTGAAGAACAGAGCAAGAGCTGTCTTAACGACGATTAACATCGGAATC 660
DB 601 CTTTAAACGTGAAGAACAGAGCAAGAGCTGTCTTAACGACGATTAACATCGGAATC 660
QY 661 AAGGTGAGGCGGAGCAATGCTCATTAACCGGTGCGTAAGCTTCGCAATCAATGCTGAG 720
DB 661 AAGGTGAGGCGGAGCAATGCTCATTAACCGGTGCGTAAGCTTCGCAATCAATGCTGAG 720
QY 721 GAGATATTTCTCTGCAACAGCAAGGAAATGCGGAGCCGGGATGGGAAAGATCAC 780
DB 721 GAGATATTTCTCTGCAACAGCAAGGAAATGCGGAGCCGGGATGGGAAAGATCAC 780
QY 781 GATTCAACACAGTGAACACCTTGTGAATGCTGATTAACGCGCAATTTATTTGCTACAA 840
DB 781 GATTCAACACAGTGAACACCTTGTGAATGCTGATTAACGCGCAATTTATTTGCTACAA 840
QY 841 CGCAGCGCAGAGGTTCGCGCAGCGCGCAACCCGCTTATGATTTGATCAAGACAGG 900
DB 841 CGCAGCGCAGAGGTTCGCGCAGCGCGCAACCCGCTTATGATTTGATCAAGACAGG 900
QY 901 TTGACGCCCATCAACCGCAAAACAGGCGTATGCTGATCAATCACTTCAGTACTG 960
DB 901 TTGACGCCCATCAACCGCAAAACAGGCGTATGCTGATCAATCACTTCAGTACTG 960

```

```

DB 901 TTGACGCCCATCAACCGCAAAACAGGCGTATGCTGATCAATCACTTCAGTACTG 960
QY 961 TTTATTTGCGGACACAGATCTAATCTGCAATCTTGCGGCGGCACTGAGCTCACTGG 1020
DB 961 TTTATTTGCGGACACAGATCTAATCTGCAATCTTGCGGCGGCACTGAGCTCACTGG 1020
QY 1021 ACCGTTCCCGGTACAGCGGATTAACAGCGCGGAGTGTGAACCTGATTTGAAACGCTGG 1080
DB 1021 ACCGTTCCCGGTACAGCGGATTAACAGCGCGGAGTGTGAACCTGATTTGAAACGCTGG 1080
QY 1081 CGTGGCTAAGCGATTAACAGCGAGTGAATTCAGTTTCGCTGCTTCAGACTTTACAG 1140
DB 1081 CGTGGCTAAGCGATTAACAGCGAGTGAATTCAGTTTCGCTGCTTCAGACTTTACAG 1140
QY 1141 CAGATGCGGTATTAACCGCGCTGTCTAATTAATAGCCGCCCGGAGAGGTGAACCTGACC 1200
DB 1141 CAGATGCGGTATTAACCGCGCTGTCTAATTAATAGCCGCCCGGAGAGGTGAACCTGACC 1200
QY 1201 CTGGCAGATGTAAGAGGAAATGCGCAGGCGATGTGCTTGGCAGAGTTTACGCA 1260
DB 1201 CTGGCAGATGTAAGAGGAAATGCGCAGGCGATGTGCTTGGCAGAGTTTACGCA 1260
QY 1261 ATCGTAATGAAGCAGCATACCGGCGTGCAGTTTGAATCTCATCAGCATCAC 1320
DB 1261 ATCGTAATGAAGCAGCATACCGGCGTGCAGTTTGAATCTCATCAGCATCAC 1320
QY 1321 TAA 1323
DB 1321 TAA 1323

RESULT 8
AD050303
ID AD050303 standard; DNA; 1323 BP.
XX
AC AD050303;
XX
XX 29-JUL-2004 (first entry)
XX
DB Escherichia coli w phytase 875PH2 mutant DNA.
XX
XX Phytase; bacteria; thermal tolerance; protease stability; foodstuff;
XX phytate; animal feed; fish feed; dough; baking; mutant; gene; ds.
XX
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT 1..1323
FT /tag= a
FT /product= "Mutant phytase enzyme"
FT 214..216
FT /tag= b
FT /note= "Encodes Arg"
FT mutation
FT /tag= c
FT replace(527,A)
FT mutation
FT /tag= d
FT replace(893,T)
FT mutation
FT /tag= e
FT replace(895,G)
FT /tag= f
XX
XX US2004091968-A1.
XX
XX 13-MAY-2004.
XX
XX 20-JUN-2003; 2003US-00601319.
XX
XX 13-AUG-1997; 97US-00910798.
XX
XX 01-MAR-1999; 99US-00259214.
XX
XX 13-APR-1999; 99US-00291931.
XX
XX 25-MAY-1999; 99US-00318528.
XX
XX 25-MAY-2000; 2000US-00580515.

```

PR 24-MAY-2001; 2001US-00866379.
XX (SHOR/) SHORT J M.
PA (KRET/) KRETZ K.
PA (GRAY/) GRAY K A.
PA (BART/) BARTON N R.
PA (GAR/) GARRETT J B.
PA (ODON/) O'DONOGHUE E.
PA (MATH/) MATHER E J.
XX
PI Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E,
PI Mather EJ;
XX
DR MPI: 2004-374952/35.
DR P-PSDB: ADO50292, ADO50304.
XX
PT Producing phytase, involves providing nucleic acid derived from bacteria
PT encoding polypeptide having phytase activity, and expressing nucleic acid
PT in yeast.
XX
PS Disclosure; Page; 74pp; English.
XX
CC The invention relates to a method for producing phytase that involves
CC providing a nucleic acid encoding phytase derived from a bacteria, and
CC expressing the nucleic acid in a yeast under conditions that allow
CC expression of the enzyme in the yeast. The invention also relates to
CC modified phytase enzyme which has improved thermal tolerance and protease
CC stability at low pH. The phytase enzyme can be used in foodstuffs to
CC improve the feeding value of phytate rich ingredients, and in diet of
CC numerous animals including mammals, fowls and fishes, commercially
CC significant avian species such as pigs, goats, laboratory rodents, commercially
CC commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC etc., in dough making and baking, in dietary aids for animals. The method
CC provides easy manufacture of the active ingredient loaded biocompatible
CC composition, higher yields and loading efficiency. The phytase
CC incorporation, in to the dietary aid is safe for animals. The present
CC sequence is *Escherichia coli* W phytase mutant DNA. Note: The present
CC sequence is not shown in the specification but has been derived from
CC *Escherichia coli* B phytase DNA ADO50291.
XX
SQ Sequence 1323 BP; 324 A; 353 C; 356 G; 289 T; 0 U; 1 Other;
Query Match 99.5%; Score 1315.6; DB 12; Length 1323;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1319; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGAAGCGATCTTAATCCATTTTATCTTCTGATTCGGTTAACCCCGCAATCTGCA 60
DB 1 ATGAAGCGATCTTAATCCATTTTATCTTCTGATTCGGTTAACCCCGCAATCTGCA 60
QY 61 TTCGCTCAGAGTGAAGCCGAGAGCTGAAAGCTGAAAGTGTGTGATTTGTCTCATGTGT 120
DB 61 TTCGCTCAGAGTGAAGCCGAGAGCTGAAAGTGTGTGATTTGTCTCATGTGTGT 120
QY 121 GTGCGGTCTCAACCAAGGCGACGCAACTGATGACAGATGTCAACCCGACGATGGCCA 180
DB 121 GTGCGGTCTCAACCAAGGCGACGCAACTGATGACAGATGTCAACCCGACGATGGCCA 180
QY 181 ACCTGGCCGGTAAACTGGGTTGGCTGACACCGCGGAGTGTGAGTAAATGCTATCTC 240
DB 181 ACCTGGCCGGTAAACTGGGTTGGCTGACACCGCGGAGTGTGAGTAAATGCTATCTC 240
QY 241 GGAACATTAACCAACGCGAGCTGTGTGAGCGGACGATTGCTGGCGAAAAAGGCTGCCG 300
DB 241 GGAACATTAACCAACGCGAGCTGTGTGAGCGGACGATTGCTGGCGAAAAAGGCTGCCG 300
QY 301 CAGTCTGAGTCAAGTGTGATTTGCTGATGTGACGAGCGTAAACCGGTAAACAGGCGAA 360
DB 301 CAGTCTGAGTCAAGTGTGATTTGCTGATGTGACGAGCGTAAACCGGTAAACAGGCGAA 360
QY 361 GCCTTCGCCCGCGGGCTGACCTGACTGTGCAATTAACCGTACATCCAGGCAAGTACG 420

DB 361 GCCTTCGCCCGCGGGCTGACCTGACTGTGCAATTAACCGTACATCCAGGCAAGTACG 420
QY 421 TCAGTCCCGATTCGGTATTTAATCTCTAATAAATCGGCGTTTGCACTGATTAACGCG 480
DB 421 TCAGTCCCGATTCGGTATTTAATCTCTAATAAATCGGCGTTTGCACTGATTAACGCG 480
QY 481 AACGTGACTGACGAGATCTTCAGCAGGAGCAGAGAGGTCATTTGCTGACTTTACCGGGCAT 540
DB 481 AACGTGACTGACGAGATCTTCAGCAGGAGCAGAGAGGTCATTTGCTGACTTTACCGGGCAT 540
QY 541 CGGCAACGCGGTTTCCGGAATCGGAACGCGTCTTAATTTTCCGCAATCAACTGTGTC 600
DB 541 CGGCAACGCGGTTTCCGGAATCGGAACGCGTCTTAATTTTCCGCAATCAACTGTGTC 600
QY 601 CTTAAAGTGTGAACCAAGGAGGAGTCTTAATCAAGGAGTCAATCAATCAATCAATCAATCA 660
DB 601 CTTAAAGTGTGAACCAAGGAGGAGTCTTAATCAAGGAGTCAATCAATCAATCAATCAATCA 660
QY 661 AAGGTGAGCGCGGCAATGTCTCATTTAACCGGTGCGTAAAGCTTCGATCAATGCTGACG 720
DB 661 AAGGTGAGCGCGGCAATGTCTCATTTAACCGGTGCGTAAAGCTTCGATCAATGCTGACG 720
QY 721 GAGATATTTCTCTGCAACAGCAAGGAAATCCCGAGCCGGGCTGGGAAAGATCAC 780
DB 721 GAGATATTTCTCTGCAACAGCAAGGAAATCCCGAGCCGGGCTGGGAAAGATCAC 780
QY 781 GATTCAACGAGTGAACACCTGTGAATTTGATTAACGCGCAATTTATTTGCTACAA 840
DB 781 GATTCAACGAGTGAACACCTGTGAATTTGATTAACGCGCAATTTATTTGCTACAA 840
QY 841 CGCAGCGCAGAGGTTGCGCGGAGCGCGCACCCTGTTATTTGATTTGATTTGATTTGATTTG 900
DB 841 CGCAGCGCAGAGGTTGCGCGGAGCGCGCACCCTGTTATTTGATTTGATTTGATTTGATTTG 900
QY 901 TTGACGCGGCTTCCACCGCAAAACAGGCGTATGTGTGATTAACCACTTCAGTACTG 960
DB 901 TTGACGCGGCTTCCACCGCAAAACAGGCGTATGTGTGATTAACCACTTCAGTACTG 960
QY 961 TTTATTTGCGGAGCAAGTACTTAATCTGCGCAATCTCGGCGGCGGAGTGAAGTCACTGG 1020
DB 961 TTTATTTGCGGAGCAAGTACTTAATCTGCGCAATCTCGGCGGCGGAGTGAAGTCACTGG 1020
QY 1021 ACGCTTCCCGTCAAGCGCGATTAACAGCGCGCAGGTGTGTAATCTGTGTTGAACGCTGG 1080
DB 1021 ACGCTTCCCGTCAAGCGCGATTAACAGCGCGCAGGTGTGTAATCTGTGTTGAACGCTGG 1080
QY 1081 CGTGGCTTAAGCGATTAACAGCGCGATGATTTGAGTTTCTGCTTCCAGACTTTACAG 1140
DB 1081 CGTGGCTTAAGCGATTAACAGCGCGATGATTTGAGTTTCTGCTTCCAGACTTTACAG 1140
QY 1141 CAGATGCGTGAATAAAGCGCGCTGTCAATTAATAGCGCGCGGAGAGGTGAACCTGACC 1200
DB 1141 CAGATGCGTGAATAAAGCGCGCTGTCAATTAATAGCGCGCGGAGAGGTGAACCTGACC 1200
QY 1201 CTGGCAGGATGTGAAGAGCGAAATGCGAGGCAATGTGTTGTTGCGAGGTTTTCGCA 1260
DB 1201 CTGGCAGGATGTGAAGAGCGAAATGCGAGGCAATGTGTTGTTGCGAGGTTTTCGCA 1260
QY 1261 ATCGTGAATGAAGCAGCATACCGGCTGCTGATTTAGATTTCAATCAATCAATCAATCA 1320
DB 1261 ATCGTGAATGAAGCAGCATACCGGCTGCTGATTTAGATTTCAATCAATCAATCAATCA 1320
QY 1321 TAA 1323
DB 1321 TAA 1323
RESULT 9
AAD25463
ID AAD25463 standard; DNA; 1901 BP.
XX
AC AAD25463;
XX

Oy		1261 ATGCTGAATGAACGCAAGCATPACCGGCGTGCAGTTTGAACT 1300
Dd		1448 ATCGTGAATGAACGCAACTAACCCGCCGTGCAGTTTGTAACT 1487
RESULT 10		
ID	ADA19449	standard; DNA; 1901 BP.
XX AC	ADA19449;	
XX AD	20-NOV-2003	(first entry)
XX DE	E. coli K12 appa gene encoding phytase.	
XX KW	Phytase; ds; gene; phytate; appa gene; animal feed; inorganic phosphate; digestion enhancement; transgenic; thermal tolerance; protease stability.	
XX OS	Escherichia coli; strain K12.	
XX FH	Key	Location/Qualifiers
FH CDS		188..1486
FT FT	/tag= a	
FT FT	/product= "phytase"	
PN XX	US2002136754-A1.	
PX PD	26-SEP-2002.	
PX PF	24-MAY-2001; 2001US-00866379.	
PR PR	13-AUG-1997; 97US-00910798. 01-MAR-1999; 99US-00259214. 13-APR-1999; 99US-00291931. 25-MAY-1999; 99US-00318528. 25-MAY-2000; 2000US-00580515.	
PA PA	(SHOR/) SHORT J M.	
PA PA	(KRETZ) KRETS R A.	
PA PA	(GRAY/) GRAY K A.	
PA PA	(BART/) BARTON N R.	
PA PA	(GARRE/) GARRETT J B.	
PA PA	(DONO/) O' DONOGHUE E.	
PI PI	(MATH/) MATHEW E J.	
PT PT	Short JM, Kreetz KA, Gray KA, Barton NR, Garrett JB; O' Donoghue E, Matchur EJ; MPI; 2003-040002/03. P-PSTDS; ADA19450.	
DR DR	Isolated Escherichia coli polynucleotide encoding a modified phytase enzyme, useful in the production of animal feed, for improving the nutritional value of phytate-containing foodstuff and for enhancing digestion in humans and animals.	
XX PS	Claim 42; Fig 7; 62pp; English.	
XX CC	The invention relates to an isolated Escherichia coli polynucleotide encoding a phytase enzyme appearing as ADA19450 and having amino acids modification from W68E, Q84W, A95F, K97C, S168E, R180I, N282C or Y277D. Also included the E. coli appa gene ADA19449 (or an oligonucleotide derived from it) or its mutant sequence ADA19452, expression vectors, host cells, a method of improving nutritional value of a phytate- containing foodstuff by contacting the phytate-containing foodstuff with a pure phytase enzyme mutated phytase (where the phytase enzyme catalyzes the liberation of inorganic phosphate from the phytate in the phytate- containing foodstuff), a method to produce an animal feed containing a microbial phytase (comprising culturing the plant cell, plant part or plant under conditions where the nucleotide sequence is expressed and converting the plant cells, plants or plant into a composition for animal feed), a feed composition for animals (comprising the plant seeds, plant cells, plant parts or plants in admixture with a phytate-containing	

	Query Match	97.4% Best Local Similarity	Score 1287.8; 99.5% Matches 1293;	DB 8; Pred. No. 0; Conservative 0;	Mismatches 7; Indels 0; Gaps 0;	Length 1901; 1901 1901
CC	foodstuff, a method to treat a human or an animal able to benefit from					
CC	digestive enhancement by the activity of an exogenous phytase enzymes					
CC	comprising administering to the human or animal the plant seed, plant					
CC	cells, plant parts or plants of a transgenic plant which is modified to					
CC	contain an expression system which expresses a nucleotide sequence					
CC	encoding a phytase enzyme, a transgenic non-human organism whose genome					
CC	comprising a heterologous nucleic acid sequence encoding a polypeptide					
CC	having phytase activity. The phytase enzyme is useful for improving the					
CC	nutritional value of phytate-containing foodstuff, in the production of					
CC	animal feed and for enhancing digestion in humans and animals. The					
CC	invented method improves thermal tolerance and protease stability. It					
CC	also improves the feeding value of phytate rich ingredients. The present					
CC	sequence represents the E. coli K12 <i>apnA</i> gene encoding wild-type phytase.					
XX	Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 0 U; 1 Other;					
QY	1 ATGAAACGATCTTAAATCCATTTTATCTTTCTGATTCGGTAAACCCGCAATCTGCA	60				
DB	188 ATGAAACGATCTTAAATCCATTTTATCTTTCTGATTCGGTAAACCCGCAATCTGCA	247				
QY	61 TTGCTCTAGATGAGCCGGAGCTGAAAGTGTGTATTTGTCACTCCTCATGCT	120				
DB	248 TTGCTCTAGATGAGCCGGAGCTGAAAGTGTGTATTTGTCACTCCTCATGCT	307				
QY	121 GTGGGTCTCAACCAAGGCCACGCAACTGATGCAAGATGTCAACCCAGACGATGGCCA	180				
DB	308 GTGGGTCTCAACCAAGGCCACGCAACTGATGCAAGATGTCAACCCAGACGATGGCCA	367				
QY	181 AACTGGCCGGTAAACTGGGTGGTGACACCGCGAGGTGTGAGCTAATGCTTATCTC	240				
DB	368 AACTGGCCGGTAAACTGGGTGGTGACACCGCGAGGTGTGAGCTAATGCTTATCTC	427				
QY	241 GGAATTAACCAAGCCCAAGCTGTGTGACCGACGAGATTTGCTCGGAAAAAGGCTCCCG	300				
DB	428 GGAATTAACCAAGCCCAAGCTGTGTGACCGACGAGATTTGCTCGGAAAAAGGCTCCCG	487				
QY	301 CAGCTGTGGTCAAGGTGCGATTTATGCTGATGTGCAAGAGGTATCCGTTAAACAGGCGAA	360				
DB	488 CAGCTGTGGTCAAGGTGCGATTTATGCTGATGTGCAAGAGGTATCCGTTAAACAGGCGAA	547				
QY	361 GCTTTTCGCGCGGGCTGCGACCTGACTGTGCAATACCTGACATACCCAGCAGATACG	420				
DB	548 GCTTTTCGCGCGGGCTGCGACCTGACTGTGCAATACCTGACATACCCAGCAGATACG	607				
QY	421 TCCAGTCCCGATCCGTTATTTATTCCTCTTAAAACTGGCTTTGCCACTGTGATTAACGCG	480				
DB	608 TCCAGTCCCGATCCGTTATTTATTCCTCTTAAAACTGGCTTTGCCACTGTGATTAACGCG	667				
QY	481 AAGGTGACTGACGCGATCCTCAGACGGGCGAGAGGGCTCAATTGCTGACTTTACCGGGCAT	540				
DB	668 AAGGTGACTGACGCGATCCTCAGACGGGCGAGAGGGCTCAATTGCTGACTTTACCGGGCAT	727				
QY	541 CGGCAAAACGCGGCTTTCCGCAATCTGAAACGGGCTCTTAAATTTTCCGCAATCAAACTTGTGC	600				
DB	728 CGGCAAAACGCGGCTTTCCGCAATCTGAAACGGGCTCTTAAATTTTCCGCAATCAAACTTGTGC	787				
QY	601 CTTAAACGTGAGAAACGAGACGAAAGCTGTTCAATTAACGAGGCATTAACATGAGAACTC	660				
DB	788 CTTAAACGTGAGAAACGAGACGAAAGCTGTTCAATTAACGAGGCATTAACATGAGAACTC	847				
QY	661 AAGGTGAGCCCGCAATGTTCTAATTAACGGGTGCGGTAGCCTCGCATCAATGCTGACG	720				
DB	848 AAGGTGAGCCCGCAATGTTCTAATTAACGGGTGCGGTAGCCTCGCATCAATGCTGACG	907				
QY	721 GAGATATTTTCTCTGCAACAAGCAACGAGGAATCCGAGCCGGGGTGGAGAAAGATCAACC	780				
DB	908 GAGATATTTTCTCTGCAACAAGCAACGAGGAATCCGAGCCGGGGTGGAGAAAGATCAACC	967				
QY	781 GATTCAACCAAGTGAACCACTTGTCTAATGTTTGATTAACGCGCAATTTTATTTTGCTACAA	840				

```

Db      968 GATTGACACAGAGGAACCTTGTAAAGTTGCATACCGCAATTTTATTTGCTACAA 1027
Qy      841 CGACGCCGAGAGTTGTCGCCGAGCGCGCCGCAATTTGATTTGATGATGAGG 900
Db      1028 CGACGCCGAGAGTTGTCGCCGAGCGCGCCGCAATTTGATTTGATGATGAGG 1087
Qy      901 TTGAGCGCCGATCCACCGCAAAAGGGGTATGATGATGATGATGATGATGATGATG 960
Db      1088 TTGAGCGCCGATCCACCGCAAAAGGGGTATGATGATGATGATGATGATGATGATG 1147
Qy      961 TTTATTTGCGGACACGATATCTATCTGCAAAATCTGCGCGCGCATGAGCTCAACTG 1020
Db      1148 TTTATTTGCGGACACGATATCTATCTGCAAAATCTGCGCGCGCATGAGCTCAACTG 1207
Qy      1021 AGCGTTCCCGGTACCGCGATTAACAGCGCGCGCGAGTGTGAACTGTGTTGAAACGCTG 1080
Db      1208 AGCGTTCCCGGTACCGCGATTAACAGCGCGCGCGAGTGTGAACTGTGTTGAAACGCTG 1267
Qy      1081 CGTCCGCTAAGGATTAACAGCGCGCGCGAGTGTGATTTCCGTTCTTCCAGACTTTACG 1140
Db      1268 CGTCCGCTAAGGATTAACAGCGCGCGCGAGTGTGATTTCCGTTCTTCCAGACTTTACG 1327
Qy      1141 CAGATGCGTGAATAAAGCGCGCTGTCAATTAATACGCGCGCGAGGAGGAACTGACC 1200
Db      1328 CAGATGCGTGAATAAAGCGCGCTGTCAATTAATACGCGCGCGAGGAGGAACTGACC 1387
Qy      1201 CTGGCAGATGTGAAGAGCGAAATGCGAGCGGATGTTCGTTGCGAGGTTTATGCA 1260
Db      1388 CTGGCAGATGTGAAGAGCGAAATGCGAGCGGATGTTCGTTGCGAGGTTTATGCA 1447
Qy      1261 ATCGTGAATGAAGCAGCATACCGCGGTGCACTTTAGAT 1300
Db      1448 ATCGTGAATGAAGCAGCATACCGCGGTGCACTTTAGAT 1487

RESULT 11
ADC87744
ID      ADC87744 standard, DNA, 1901 BP.
XX
AC      ADC87744;
XX
DT      01-JAN-2004 (first entry)
XX
DE      DNA encoding Escherichia coli appa phytase.
XX
KM      Phytase; food supplement; enzyme delivery matrix; soybean meal;
XX      myo-inositol-hexaphosphate; inositol; inorganic phosphate;
XX      thermocolerant; feed value; digestion; gene; ds; appa.
OS      Escherichia coli.
XX
XX      Key      Location/Qualifiers
FH      CDS      188..1486
FT      /tag=a
FT      /product="appa phytase"
FT      /transl_except=(pos:401..403,aa:Arg)
XX
XX      US2003103958-A1.
XX
PD      05-JUN-2003.
XX
PF      24-MAY-2002; 2002US-00156660.
XX
XX      13-AUG-1997; 97US-00910798.
XX      01-MAR-1999; 99US-00259214.
XX      13-APR-1999; 99US-00291931.
XX      25-MAY-1999; 99US-00318528.
XX      25-MAY-2000; 2000US-00580515.
XX      24-MAY-2001; 2001US-00866379.
XX
PA      (DIVE-) DIVERSA CORP.

```

```

XX      Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI      Machur EJ;
XX      WPI; 2003-787039/74.
DR      P-PSDB; ADC87745.
XX
PT      New nucleic acid encoding a polypeptide having phytase activity, useful
PT      in improving the feeding value of phytate rich ingredients or as an aid
PT      in phytate digestion.
XX
PS      Example 1; SEQ ID NO 3; 113pp; English.
XX
CC      The invention discloses a new isolated or recombinant nucleic acid which
CC      encodes a polypeptide having a phytase activity. Also claimed is a
CC      nucleic acid probe, an amplification primer sequence pair, an expression
CC      cassette comprising the nucleic acid, a vector comprising the nucleic
CC      acid, a transgenic non-human animal or plant, or its seed, comprising the
CC      nucleic acid, an antisense oligonucleotide, inhibiting the translation of
CC      a phytase message in a cell, a heterodimer comprising the polypeptide and
CC      a second domain, an array comprising immobilised polypeptide or nucleic
CC      acid, a hydridoma comprising an antibody that specifically binds to the
CC      polypeptide, a food supplement for an animal, an edible enzyme delivery
CC      matrix, an edible pellet comprising a granule edible carrier and the
CC      polypeptide, a feed composition, a soybean meal, isolating or identifying
CC      the polypeptide, making an anti-phytase antibody, producing a recombinant
CC      polypeptide, determining whether a compound binds to the polypeptide,
CC      identifying a modulator, whole cell engineering of new or modified
CC      phenotypes by using real-time metabolic flux analysis, increasing
CC      thermotolerance or thermostability of the phytase polypeptide, increasing
CC      the resistance of the phytase polypeptide to enzymatic inactivation in a
CC      digestive system and processing of corn and sorghum kernels. The phytase
CC      activity comprises catalysis of phytate (myo-inositol-hexaphosphate) to
CC      inositol and inorganic phosphate and is thermocolerant. The nucleic acid
CC      is useful in improving the feeding value of phytate rich ingredients or
CC      as an aid in phytate digestion. The sequence presented is the DNA
CC      encoding the Escherichia coli appa phytase.
XX
SQ      Sequence 1901 BP; 474 A; 499 G; 428 T; 0 U; 1 Other;
XX
Query Match      97.4%; Score 1287.8; DB 10; Length 1901;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1293; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy      1 ATGAAGCGATCTTAATCCATTTTATCTCTTGTGATTCGTTAACCCCGCAATCTGCA 60
Db      188 ATGAAGCGATCTTAATCCATTTTATCTCTTGTGATTCGTTAACCCCGCAATCTGCA 247
Qy      61 TTGCGTCAAGTAGACCCGAGCTGAAGCTGGAAGTGTGTGATTTGCTCATGTGT 120
Db      248 TTGCGTCAAGTAGACCCGAGCTGAAGTGTGTGATTTGCTCATGTGT 307
Qy      121 GTGGGTGCTCCAAACCAAGGCCACGCACTGATGAGATGTCACCCGAGGAGCATGGCA 180
Db      308 GTGGGTGCTCCAAACCAAGGCCACGCACTGATGAGATGTCACCCGAGGAGCATGGCA 367
Qy      181 ACCTGCGCGGTAAACCTGCGTTGCTGACACCGCGAGTGTGAGCTAATCGCTATCTC 240
Db      368 ACCTGCGCGGTAAACCTGCGTTGCTGACACCGCGAGTGTGAGCTAATCGCTATCTC 427
Qy      241 GGACATTACCAACGCGACGCTGTGTAGCCGAGATTGCTGGGCAAAAGGGCTGCCCC 300
Db      428 GGACATTACCAACGCGACGCTGTGTAGCCGAGATTGCTGGGCAAAAGGGCTGCCCC 487
Qy      301 CAGCTGTGACGGTGTGCAATTAATGCTGATGTCAGCGAGCGTAACCGGTAAACAGCGCA 360
Db      488 CAGCTGTGACGGTGTGCAATTAATGCTGATGTCAGCGAGCGTAACCGGTAAACAGCGCA 547
Qy      488 CAGCTGTGACGGTGTGCAATTAATGCTGATGTCAGCGAGCGTAACCGGTAAACAGCGCA 547
Qy      361 GCCTTGCGCGCGCGGCTGTGCACTGTGCAATTAACGTAATACCCAGGAGATACG 420
Db      548 GCCTTGCGCGCGCGGCTGTGCACTGTGCAATTAACGTAATACCCAGGAGATACG 607
Qy      421 TCAGTCCCGATCCGTATTATTATCTCTTAATAAAGTGGCGTTTGCAACTGATTAACGG 480

```

DB 608 TCCAGTCCCGATCCCTTATTATATCTCTAATAAACTGGGCTTTCGCAATGATATACCG 667
 QY 481 AACTGACTGACGCAATCTCTGACAGGGGCAAGGGTCAATGCTGACTTTACCGGGCAT 540
 DB 668 AACGTGACTGACGCAATCTCTGACAGGGGCAAGGGTCAATGCTGACTTTACCGGGCAT 727
 QY 541 CGGCAAAAGGGGCTTTCGGAATGGAACGGGGCTTATTTTCCCAATCAACTGAGC 600
 DB 728 CGGCAAAAGGGGCTTTCGGAATGGAACGGGGCTTATTTTCCCAATCAACTGAGC 787
 QY 601 CTTAAACGTGAGAAAACAGGACGAAAGCTGTTCAATTAACGACGCAATTAACATCGAATC 660
 DB 788 CTTAAACGTGAGAAAACAGGACGAAAGCTGTTCAATTAACGACGCAATTAACATCGAATC 847
 QY 661 AAGTGAGCGCGCAATGCTCTCAATTAACCGGTGCGTAAAGCTTCGCAATGCTGACG 720
 DB 848 AAGTGAGCGCGCAATGCTCTCAATTAACCGGTGCGTAAAGCTTCGCAATGCTGACG 907
 QY 721 GAGATATTTCTCTGCAACAAACAGGAAATGCGGAGCGGGGGTGGGAAAGATCAAC 780
 DB 908 GAGATATTTCTCTGCAACAAACAGGAAATGCGGAGCGGGGGTGGGAAAGATCAAC 967
 QY 781 GATTCAACACAGTGAACAACCTTGCTAAGTTGCAATTAACGCGCAATTTTATTGCTACAA 840
 DB 968 GATTCAACACAGTGAACAACCTTGCTAAGTTGCAATTAACGCGCAATTTTATTGCTACAA 1027
 QY 841 CGCAGCCCAAGAGTTGCGCGGACCGCGCAACCCCGTTATTTGATGATCAATGCAACG 900
 DB 1028 CGCAGCCCAAGAGTTGCGCGGACCGCGCAACCCCGTTATTTGATGATCAATGCAACG 1087
 QY 901 TTGAGCGCCCAATCCACCGGAAAACAGGCGTAAAGTGAACATTAACCACTTCAATACG 960
 DB 1088 TTGAGCGCCCAATCCACCGGAAAACAGGCGTAAAGTGAACATTAACCACTTCAATACG 1147
 QY 961 TTTATTTGCGGACAGATATCTAATCTGCAATTAATCTCGGCGGCACTGAGCTCAATCG 1020
 DB 1148 TTTATTTGCGGACAGATATCTAATCTGCAATTAATCTCGGCGGCACTGAGCTCAATCG 1207
 QY 1021 ACGCTTCCCGGTACGCGGATTAACAGCCCGCAAGGTGTGAATCTGATTTGAACGCTGG 1080
 DB 1208 ACGCTTCCCGGTACGCGGATTAACAGCCCGCAAGGTGTGAATCTGATTTGAACGCTGG 1267
 QY 1081 CSTGCGCTAAGGATTAACAGCGGATTAAGTTGCTGCTGCTTCCAGCTTTACG 1140
 DB 1268 CSTGCGCTAAGGATTAACAGCGGATTAAGTTGCTGCTGCTTCCAGCTTTACG 1327
 QY 1141 CAGATGCGTAAATAACGCGCTGCTCAATTAATACGCGCGCGGAGAGGTGAACCTGACC 1200
 DB 1328 CAGATGCGTAAATAACGCGCTGCTCAATTAATACGCGCGCGGAGAGGTGAACCTGACC 1387
 QY 1201 CTGCGAGATGTAAGAGCGAAATGCGGAGGAGCATGTTGCTGCGAGGTTTACGCA 1260
 DB 1388 CTGCGAGATGTAAGAGCGAAATGCGGAGGAGCATGTTGCTGCGAGGTTTACGCA 1447
 QY 1261 ATCGTAATGAAGACGCAATACCGGCGTCAATTTGAGAT 1300
 DB 1448 ATCGTAATGAAGACGCAATACCGGCGTCAATTTGAT 1487

RESULT 12

ID AD050297 standard; DNA; 1901 BP.

AC AD050297;

DT 29-JUN-2004 (first entry)

XX Escherichia coli K12 appa phytase DNA.

KW Appa phytase; bacteria; thermal tolerance; protease stability; foodstuff;

XX phytate; animal feed; fish feed; dough; baking; gene; ds.

OS Escherichia coli; K12.
 XX Key Location/Qualifiers
 FT CDS 188..1486
 FT /tag= a
 FT /product= "Appa phytase"
 FT unsure 401..403
 FT /tag= b
 FT /note= "Encodes Arg"

US2004091968-A1.

13-MAY-2004.

20-JUN-2003; 2003US-00601319.

13-AUG-1997; 97US-00910798.

01-MAR-1999; 99US-00259214.

13-APR-1999; 99US-00291911.

25-MAY-1999; 99US-00318528.

25-MAY-2000; 2000US-00580515.

24-MAY-2001; 2001US-00866379.

(SHOR/) SHORT J M.

(KRET/) KRETZ K.

(GRAY/) GRAY K A.

(BART/) BARTON N R.

(GAR/) GARRETT J B.

(ODON/) O'DONOGHUE E.

(MATH/) MATHER E J.

Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;

Mather EJ;

WPI; 2004-374952/35.

P-Psdb; AD050298.

Producing phytase, involves providing nucleic acid derived from bacteria

encoding polypeptide having phytase activity, and expressing nucleic acid

in yeast.

Disclosure; SEQ ID NO 7; 74bp; English.

The invention relates to a method for producing phytase that involves

providing a nucleic acid encoding phytase derived from a bacteria, and

expressing the nucleic acid in a yeast under conditions that allow

expression of the enzyme in the yeast. The invention also relates to

modified phytase enzyme which has improved thermal tolerance and protease

stability at low pH. The phytase enzyme can be used in foodstuffs to

improve the feeding value of phytate rich ingredients, and in diet of

numerous animals including mammals, fowls and fishes, commercially

significant mammals such as pigs, goats, laboratory rodents, commercially

significant avian species such as chicken, ducks, doves, parrot, etc.,

commercially farmed fish such as guppy, zebrafish, mollly, swordtail,

etc., in dough making and baking, in dietary aids for animals. The method

provides easy manufacture of the active ingredient loaded biocompatible

composition, higher yields and loading efficiency. The phytase

incorporated in to the dietary aid is safe for animals. The present

sequence is Escherichia coli K12 appa phytase DNA.

Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 0 U; 1 Other;

Query Match 97.4%; Score 1287.8; DB 12; Length 1901;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1293; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATGAAGGATCTAATCCATTTTATCTCTGATTTCCGTTAACCCGCAATCTGCA 60
 DB 188 ATGAAGGATCTAATCCATTTTATCTCTGATTTCCGTTAACCCGCAATCTGCA 247
 QY 61 TTGCTCAGAGTGAAGCCGAGCTGAAGCTGAAGAGTGTGATGTTGAGTGTGATGTT 120
 DB 248 TTGCTCAGAGTGAAGCCGAGCTGAAGCTGAAGAGTGTGATGTTGAGTGTGATGTT 307

QY 121 GTCGTCCTCAACCAAGGCCAAGCACTGATGAGATGTCACCCCAAGCGCATGCCA 180
 DB 308 GTGCGTCTCAACCAAGGCCAAGCACTGATGAGATGTCACCCCAAGCGCATGCCA 367
 QY 181 ACCTGGCCGGTAAACTGGGTTGGCTGACACCGCGNGTGTGAGCTAATCGCTATCTC 240
 DB 368 ACCTGGCCGGTAAACTGGGTTGGCTGACACCGCGNGTGTGAGCTAATCGCTATCTC 427
 QY 241 GGCATTACCAAGCCAGCGTCTGTGATGCCAGCATTTGCTGGCGAAAAAGGCTGGCCG 300
 DB 428 GGCATTACCAAGCCAGCGTCTGTGATGCCAGCATTTGCTGGCGAAAAAGGCTGGCCG 487
 QY 301 CAGTCTGTCAAGTCCGAGTATTTAGTGTGATGCCAGCATTTGCTGGCGAAAAAGGCTGGCCG 360
 DB 488 CAGTCTGTCAAGTCCGAGTATTTAGTGTGATGCCAGCATTTGCTGGCGAAAAAGGCTGGCCG 547
 QY 361 GCGTTGGCCGCGGGGCTGGCACTGTGCAATACCTGATACATCCAGGAGATACG 420
 DB 548 GCGTTGGCCGCGGGGCTGGCACTGTGCAATACCTGATACATCCAGGAGATACG 607
 QY 421 TCCAGTCCGAGTCCGTTATTTAATCTCTAATAACTGGCGTTGGCAACTGATTAACGCG 480
 DB 608 TCCAGTCCGAGTCCGTTATTTAATCTCTAATAACTGGCGTTGGCAACTGATTAACGCG 667
 QY 481 AACGTACTGAAGCGATCCCTGAGCAAGGAGGATCAATTTGCTGATCCGGGCAAT 540
 DB 668 AACGTACTGAAGCGATCCCTGAGCAAGGAGGATCAATTTGCTGATCCGGGCAAT 727
 QY 541 CGGCAACCGCGCTTTGCGCGAATGGAACGGGTCTTAATTTTCCGCAATCAAACTTGTGC 600
 DB 728 CGGCAACCGCGCTTTGCGCGAATGGAACGGGTCTTAATTTTCCGCAATCAAACTTGTGC 787
 QY 601 CTTTAAAGTGAAGAAAGAGACGAAGCTGTTCAATTAACGAGCATTTACATCGAACTC 660
 DB 788 CTTTAAAGTGAAGAAAGAGACGAAGCTGTTCAATTAACGAGCATTTACATCGAACTC 847
 QY 661 AAGGTGAGGGCGGAGCAATGTTCTCATTAACGGTGGCTGAAGCTCCGATCAATGCTGAGC 720
 DB 848 AAGGTGAGGGCGGAGCAATGTTCTCATTAACGGTGGCTGAAGCTCCGATCAATGCTGAGC 907
 QY 721 GAGATATTTCTCTGTGCAAGACAGGGAATGCCGAGCCGGGGTGGGGAAGATCACCC 780
 DB 908 GAGATATTTCTCTGTGCAAGACAGGGAATGCCGAGCCGGGGTGGGGAAGATCACCC 967
 QY 781 GATTCACACAGTGAACACCTTTGCTAGTTTGATTAACGCGCAATTTTATTTGCTACAA 840
 DB 968 GATTCACACAGTGAACACCTTTGCTAGTTTGATTAACGCGCAATTTTATTTGCTACAA 1027
 QY 841 CGCAGCGGAGAGGTTGCCCGGACCGCGGCAACCCGTTATTGATTTGATGAGGACGCG 900
 DB 1028 CGCAGCGGAGAGGTTGCCCGGACCGCGGCAACCCGTTATTGATTTGATGAGGACGCG 1087
 QY 901 TTGAGCGCCCATTCACCGGAAAAACAGGGGTATGATGATCAATTAACCACTTCAGTACTG 960
 DB 1088 TTGAGCGCCCATTCACCGGAAAAACAGGGGTATGATGATCAATTAACCACTTCAGTACTG 1147
 QY 961 TTTATTTGCGGACACAGTACTAATCTGGAATACTTCGGGCGGCGCATCTGAGCTCAACTG 1020
 DB 1148 TTTATTTGCGGACACAGTACTAATCTGGAATACTTCGGGCGGCGCATCTGAGCTCAACTG 1207
 QY 1021 AAGGTTCCCGGTCAAGCCGGAATTAACAGCGCGGAGGTTGGAATCTGTGTTTGAACGCTG 1080
 DB 1208 AAGGTTCCCGGTCAAGCCGGAATTAACAGCGCGGAGGTTGGAATCTGTGTTTGAACGCTG 1267
 QY 1081 CGTCGGCTAAGCGATTAACAGCCAGTGTGATTTGCTGGTCTTCAGACTTTACAG 1140
 DB 1268 CGTCGGCTAAGCGATTAACAGCCAGTGTGATTTGCTGGTCTTCAGACTTTACAG 1327
 QY 1141 CAGATGCGTGAATTAACGCGCTGTCTAATTAATACGCGCGCGGAGAGGTGAACCTGACC 1200
 DB 1328 CAGATGCGTGAATTAACGCGCTGTCTAATTAATACGCGCGCGGAGAGGTGAACCTGACC 1387

QY 1201 CTGGCAGATGTGAAGAGGAAATGCGCAGGGCATGTGTTGTTGSCAGGTTTACGCA 1260
 DB 1388 CTGGCAGATGTGAAGAGGAAATGCGCAGGGCATGTGTTGTTGSCAGGTTTACGCA 1447
 QY 1261 ATCTGTAATGAAGACGATACCGGCGTGCAGTTTGAGAT 1300
 DB 1448 ATCTGTAATGAAGACGATACCGGCGTGCAGTTTGTAAT 1487
 RESULT 13
 AED50819
 ID AED50819 standard; DNA; 1901 BP.
 XX
 AC AED50819;
 XX
 DT 29-DEC-2005 (first entry)
 XX
 DE Escherichia coli phytase appA gene.
 XX
 KW Transgenic animal; gene therapy; feedstuff; food; thermostable; ds; gene;
 KM 3-phytase; appA gene.
 OS Escherichia coli; strain K12.
 XX
 FH Key
 FT CDS location/Qualifiers
 FT 188..1486
 FT /tag= a
 FT /product= "Escherichia coli phytase protein"
 FT /transl_except= (pos:401..403,aa:Arg)
 FT /EC_number= "3.1.3.8"
 XX
 PN US2005246780-A1.
 XX
 PD 03-NOV-2005.
 XX
 PF 11-FEB-2005; 2005US-00056354.
 XX
 PR 13-AUG-1997; 97US-00910798.
 PR 01-MAR-1999; 99US-00259214.
 PR 13-APR-1999; 99US-00291931.
 PR 25-MAY-1999; 99US-00318528.
 PR 25-MAY-2000; 2000US-00580515.
 PR 24-MAY-2001; 2001US-00866379.
 PR 24-MAY-2002; 2002US-00156660.
 XX
 PA (SHOR/) SHORT J M.
 PA (KRETZ/) KRETZ K.
 PA (GRAY/) GRAY K A.
 PA (BART/) BARTON N R.
 PA (GAR/) GARRETT J B.
 PA (ODON/) ODOGHUE E.
 PA (MATH/) MATHUR E J.
 XX
 PI Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, Odoghue E;
 PI Mathur EJ;
 XX
 DR WPI; 2005-733851/75.
 DR P-PsDB; AED50820.
 DR GENBANK; M58708.
 XX
 PT New isolated or recombinant nucleic acid encoding phytase enzymes, useful
 PT as foodstuff, and for oil degumming, producing an animal feed, delivering
 PT a phytase enzyme supplement to an animal.
 XX
 PS Example 1; SEQ ID NO 3; 104dp; English.
 XX
 CC The present invention relates to a recombinant phytase (EC 3.1.2.8)
 CC protein and its encoding also nucleic acid molecule. Phytases of the
 CC invention is a thermostable protein. They are useful in foodstuffs, for
 CC oil degumming, producing an animal feed, delivering a phytase enzyme
 CC supplement to an animal, increasing the resistance of a phytase
 CC polypeptide to enzymatic inactivation in a digestive system of an animal
 CC and in processing of corn and sorghum kernels. The invention is also

CC useful in gene therapy and in production of transgenic animals. The
 CC present sequence is a *Escherichia coli* phytase appa gene. Note: The
 CC description specified in the specification for the SEQ ID NO:3 is *E. coli*
 CC phytase appa gene. However the description specified in the accession
 CC number M58708 is *E. coli* periplasmic phosphanhydride phosphohydrolase
 CC (appa) gene. Thus the description given in the specification varies with
 CC the accession number.

XX
 SQ Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 0 U; 1 Other;

Query Match 97.4%; Score 1287.8; DB 14; Length 1901;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 1293; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

QY 1 ATGAAAGGATCTTAAATCCATTTTATCTCTTGATTTCCGTTAAACCCCGAATCTGCA 60
DB 188 ATGAAAGGATCTTAAATCCATTTTATCTCTTGATTTCCGTTAAACCCCGAATCTGCA 247
QY 61 TTGCTCAGAGTGAAGCCGAGCTGAAGCTGAAAGTGTGATGTTGTCAGTGTGATGAT 120
DB 248 TTGCTCAGAGTGAAGCCGAGCTGAAGCTGAAAGTGTGATGTTGTCAGTGTGATGAT 307
QY 121 GTGCGTGTCTCAACCAAGGCCACGCACTGATGACGATGTCAACCCAGACGATGACCA 180
DB 308 GTGCGTGTCTCAACCAAGGCCACGCACTGATGACGATGTCAACCCAGACGATGACCA 367
QY 181 ACTGCGCCGGTAAACTGGGTTGGCTGACACCGCGNGTGGTGAAGTAAATGCGCTATCTC 240
DB 368 ACTGCGCCGGTAAACTGGGTTGGCTGACACCGCGNGTGGTGAAGTAAATGCGCTATCTC 427
QY 241 GGAACATTAACCAACGCGAGGCTGTGTAAGCCAGGATTTGCTGCGAAAAAGGGCTGCCCC 300
DB 428 GGAACATTAACCAACGCGAGGCTGTGTAAGCCAGGATTTGCTGCGAAAAAGGGCTGCCCC 487
QY 301 CAGTCTGTGATGAGTGCAGATTAATGCTGATGTGACGAGCGTAAACCCGTAACCAAGCGAA 360
DB 488 CAGTCTGTGATGAGTGCAGATTAATGCTGATGTGACGAGCGTAAACCCGTAACCAAGCGAA 547
QY 361 GCCTTGGCCCGGGGCTGGACCTGACCTGTGCAATTAACCGTACATTAACCGGACGATACG 420
DB 548 GCCTTGGCCCGGGGCTGGACCTGACCTGTGCAATTAACCGTACATTAACCGGACGATACG 607
QY 421 TCCAGTCCCGATCGGTTATTTAACTCTCTAAAAAACGCGGTTTGGCAATGGAATAACGG 480
DB 608 TCCAGTCCCGATCGGTTATTTAACTCTCTCTAAAAAACGCGGTTTGGCAATGGAATAACGG 667
QY 481 AACGTGACTGACGCGATCTCTGACGAGGCGAGGAGGTCAATTTGACTTTTAACCGGGCAT 540
DB 668 AACGTGACTGACGCGATCTCTGACGAGGCGAGGAGGTCAATTTGACTTTTAACCGGGCAT 727
QY 541 CGGCAAAAGGCGTTTCCGGAATGGAAGGGGTCTTAATTTTCCGCAATCAAACTTTGGC 600
DB 728 CGGCAAAAGGCGTTTCCGGAATGGAAGGGGTCTTAATTTTCCGCAATCAAACTTTGGC 787
QY 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAAGCAGGACCTTAACCATCGGAATCC 660
DB 788 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAAGCAGGACCTTAACCATCGGAATCC 847
QY 661 AAGGTGAGGCGCGCAATGTCTCATTAACCGGTGCGGTAGAGCTGCAATCAATGCTGACG 720
DB 848 AAGGTGAGGCGCGCAATGTCTCATTAACCGGTGCGGTAGAGCTGCAATCAATGCTGACG 907
QY 721 GAGATATTTCTCTGCAACAGCAACGAGGATGCGGAGCCGCGGTGGGAAAGATCAAC 780
DB 908 GAGATATTTCTCTGCAACAGCAACGAGGATGCGGAGCCGCGGTGGGAAAGATCAAC 967
QY 781 GATTACACCAATGGAACAACCTTGTCTAAGTTGTGCAATACGCCAATTTTATTTGCTACAA 840
DB 968 GATTACACCAATGGAACAACCTTGTCTAAGTTGTGCAATACGCCAATTTTATTTGCTACAA 1027
QY 841 CGCAGCGCAGAGGTTGCCGCGAGCGCGCACCCGTTATTTGATTTGATCAAGCAGCG 900
DB 1028 CGCAGCGCAGAGGTTGCCGCGAGCGCGCACCCGTTATTTGATTTGATCAAGCAGCG 1087

```

```

QY 901 TTGACGCCCATTCACCGCAAAAACAGGCGGTATGTGACATTACCACTTCACTG 960
DB 1088 TTGACGCCCATTCACCGCAAAAACAGGCGGTATGTGACATTACCACTTCACTG 1147
QY 961 TTTATTTGCCGACACGATTAATCTGCGAAATCTCGCGCGGACACTGAGACTCACTGG 1020
DB 1148 TTTATTTGCCGACACGATTAATCTGCGAAATCTCGCGCGGACACTGAGACTCACTGG 1207
QY 1021 ACGCTTCCCGGTGACCGCGATTAACACCGCCGACAGTGTGAACTGTGTTGAAACGCTGG 1080
DB 1208 ACGCTTCCCGGTGACCGCGATTAACACCGCCGACAGTGTGAACTGTGTTGAAACGCTGG 1267
QY 1081 CGTGGCTTAAGCGATTAACAGCAGTGTGATTCAGGTTTCGCTGAGCTTCCAGACTTTACAG 1140
DB 1268 CGTGGCTTAAGCGATTAACAGCAGTGTGATTCAGGTTTCGCTGAGCTTCCAGACTTTACAG 1327
QY 1141 CAGATGCGTGAATAAAGCCGCTGTCAATTAATACGCCGCCGAGAGGTGAAACTGAC 1200
DB 1328 CAGATGCGTGAATAAAGCCGCTGTCAATTAATACGCCGCCGAGAGGTGAAACTGAC 1387
QY 1201 CTGCGAGATGTGAAGACGAAATGCGCAGGCGATGTGTTCTGTGCAAGTTTAAACGCA 1260
DB 1388 CTGCGAGATGTGAAGACGAAATGCGCAGGCGATGTGTTCTGTGCAAGTTTAAACGCA 1447
QY 1261 ATCGTGAATGAAGACGATACCGGCGGTGCAAGTTGAGAT 1300
DB 1448 ATCGTGAATGAAGACGATACCGGCGGTGCAAGTTGAGAT 1487

```

RESULT 14

AE875418
 ID AE875418 standard; DNA; 1901 BP.

AC AE875418;

DT 23-FEB-2006 (first entry)

DE *E. coli* B phytase appa coding sequence SEQ ID NO: 7.

KW ds; gene; thermostable; phytase; protein stabilization; pharmaceutical;

KW fertilizer; cramp; cell culture; osteoporosis; osteopathnc; food; appa.

OS *Escherichia coli* K12.

XX Key Location/Qualifiers

FT CDS 188..1483

FT /*tag= a

FT /product= "Phytase"

FT 401..403

FT /*tag= b

FT /note= "ecodes Arg"

PN US2005281792-A1.

PD 22-DEC-2005.

PF 01-SEP-2004; 2004US-00933115.

PR 13-AUG-1997; 97US-00910798.

PR 01-MAR-1989; 99US-00259214.

PR 13-APR-1989; 99US-00291931.

PR 25-MAY-1999; 99US-00318528.

PR 25-MAY-2000; 2000US-00580515.

PR 24-MAY-2001; 2001US-00866379.

XX (SHORT) SHORT J M.

XX (KRETZ) KRETZ K A.

XX (GRAY) GRAY K A.

XX (BART) BARTON N R.

XX (GARR) GARRETT J B.

XX (ODON) ODOONOHUE E.

XX (BAUM) BAUM W.

PA (ROBE/) ROBERTSON D E.
 PA (ZORN/) ZORNER P.
 PI Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, Odonoghue B,
 PI Baum W, Robertson DE, Zorner P;
 XX MPI: 2006-055022/06.
 DR P-PSDB; ABE75419.
 XX
 PT Formulation useful as dietary supplement for treating, preventing or
 PT reversing osteoporosis and bone loss, and preventing muscle cramps,
 PT comprises one or more polypeptides having phytase activity.
 XX
 PS Claim 1; SEQ ID NO 7; 82pp; English.
 XX
 CC The present sequence is that of a polynucleotide encoding the Escherichia
 CC coli phytase appA, which was modified to produce a thermostable prease
 CC resistant phytase which constitutes the novel enzyme of the invention.
 CC The present invention relates to a novel formulation, comprising at least
 CC one polypeptide having phytase activity, where the polypeptide is a
 CC modified sequence (ABE75421) derived from the Escherichia coli K-12 appA
 CC phytase (ABE75419). The modification of the enzyme, comprising a series
 CC of mutations, improves the thermal tolerance and protease stability of
 CC the protein. The specification also claims a pharmaceutical composition,
 CC a kit; an immobilized phytase; a fertilizer or soil additive; a liquid
 CC supplement for preventing muscle cramps; a hydrating agent; a tissue
 CC culture or cell culture media; and a plant food additive, all comprising
 CC the novel phytase, and a method of reducing pollution and increasing
 CC nutrient availability in an environment or environmental sample by
 CC degrading environmental phytic acid. The novel enzyme is a dietary
 CC supplement useful for treating, preventing or reversing osteoporosis or
 CC bone loss, and preventing muscle cramps. The liquid supplement is useful
 CC for preventing muscle cramps. The method is useful for reducing pollution
 CC and increasing nutrient availability in an environment or environmental
 CC sample by degrading environmental phytic acid, where the environmental
 CC environmental sample comprises a soil or a body of water. The immobilized
 CC phytase is useful in foodstuffs for improving the feeding value of
 CC phytate rich ingredients.
 CC
 XX
 SQ Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 0 U; 1 Other;

Query Match 97.4%; Score 1287.8; DB 15; Length 1901;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1293; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATGAAGCGATCTTAATCCCATTTTATCTCTTGATTCGGTTAACCCCGAATCTGCA 60
 DB 188 ATGAAGCGATCTTAATCCCATTTTATCTCTTGATTCGGTTAACCCCGAATCTGCA 247
 QY 61 TTCGCTCAGAGTGAAGCCGGAAGCTGAAAGTGTGTGATTTCTAGTGTCTATGTT 120
 DB 248 TTCGCTCAGAGTGAAGCCGGAAGCTGAAAGTGTGTGATTTCTAGTGTCTATGTT 307
 QY 121 GTGCGTGTCCAAACGCAAGGCGCAAGCTGATGCAAGATGCAACCCGAGCATGGCCA 180
 DB 308 GTGCGTGTCCAAACGCAAGGCGCAAGCTGATGCAAGATGCAACCCGAGCATGGCCA 367
 QY 181 ACCCTGCGGTAAACTGTGGTGTGCTGACACCGCGGTGTGTGAGCTATTCCTATCTC 240
 DB 368 ACCCTGCGGTAAACTGTGGTGTGCTGACACCGCGGTGTGTGAGCTATTCCTATCTC 427
 QY 241 GGAATTAACCAACGCAAGGCTGTGTGATGCGAAGTTCTGCGCAAAAAGGCTGCCG 300
 DB 428 GGAATTAACCAACGCAAGGCTGTGTGATGCGAAGTTCTGCGCAAAAAGGCTGCCG 487
 QY 301 CAGTCTGTGTCAGGTGCGGATTAATGCTGATGTCAGACGATACCGGTAAACAGGCGCA 360
 DB 488 CAGTCTGTGTCAGGTGCGGATTAATGCTGATGTCAGACGATACCGGTAAACAGGCGCA 547
 QY 361 GCGTTTCGCGCGCGGCTGCGACCTTGAATGTGCAATTAACCGTATACCGAGCAATAGC 420
 DB 548 GCGTTTCGCGCGCGGCTGCGACCTTGAATGTGCAATTAACCGTATACCGAGCAATAGC 607

QY 421 TCAGTCCCGATCCGTTATTTATCTCTTAAAAAATCGGCTTTGCCAATGATTAACGCG 480
 DB 608 TCAGTCCCGATCCGTTATTTATCTCTTAAAAAATCGGCTTTGCCAATGATTAACGCG 667
 QY 481 AACGTGATGACGCGATCTCGACAGGCGAGAGGGTCAATTGTGACTTTTACCGGGCAT 540
 DB 668 AACGTGATGACGCGATCTCGACAGGCGAGAGGGTCAATTGTGACTTTTACCGGGCAT 727
 QY 541 CGGCAAAACGCGCTTTTCGCGAATCTGAAACGGGTGCTTAATTTTCCGCAATCAACTTGTGC 600
 DB 728 CGGCAAAACGCGCTTTTCGCGAATCTGAAACGGGTGCTTAATTTTCCGCAATCAACTTGTGC 787
 QY 601 CTTAAACGTGAAACACAGGACGAAGCTTTTCAATTAACGAGCATTTACATTCGGAATC 660
 DB 788 CTTAAACGTGAAACACAGGACGAAGCTTTTCAATTAACGAGCATTTACATTCGGAATC 847
 QY 661 AAGGTAGGCGCGCAAAATGCTCATTTAACCGGTGCGGTAAAGCTGCGATCAATGCGAGC 720
 DB 848 AAGGTAGGCGCGCAAAATGCTCATTTAACCGGTGCGGTAAAGCTGCGATCAATGCGAGC 907
 QY 721 GAGATATTTTCTCTGCAACAGACAGGGAATGCGGAGCCGCGGTGCGGAAGATCAAC 780
 DB 908 GAGATATTTTCTCTGCAACAGACAGGGAATGCGGAGCCGCGGTGCGGAAGATCAAC 967
 QY 781 GATTCAACACAGTGAACACCTTGCTAAGTTTGCAATACGCGCAATTTTATTGCTACAA 840
 DB 968 GATTCAACACAGTGAACACCTTGCTAAGTTTGCAATACGCGCAATTTTATTGCTACAA 1027
 QY 841 CGCAGCGCAAGAGTTTGCAGCGCGCGCACCCCGTTATTTGATTTGATCAATGAGAGC 900
 DB 1028 CGCAGCGCAAGAGTTTGCAGCGCGCGCACCCCGTTATTTGATTTGATCAATGAGAGC 1087
 QY 901 TTGAGCGCCCATCAACCGCAAAAACAGGCGTATGTGTGACATTAACCACTTCAGTACTG 960
 DB 1088 TTGAGCGCCCATCAACCGCAAAAACAGGCGTATGTGTGACATTAACCACTTCAGTACTG 1147
 QY 961 TTTTATTTGCGGACACAGATTAATCTATGCGAAATCTCGCGCGGCACTGAGCTCAACTGG 1020
 DB 1148 TTTTATTTGCGGACACAGATTAATCTATGCGAAATCTCGCGCGGCACTGAGCTCAACTGG 1207
 QY 1021 ACGCTTCCCGGTACCGCGGATTAACAGCGCGCGCAAGGTGTGAATCTGTGTTGAAGCTGG 1080
 DB 1208 ACGCTTCCCGGTACCGCGGATTAACAGCGCGCGCAAGGTGTGAATCTGTGTTGAAGCTGG 1267
 QY 1081 CGTCCGCTTAAGGATTAACAGCGGATGATTAAGTTTGTGCTGCTTCCAGACTTTACAG 1140
 DB 1268 CGTCCGCTTAAGGATTAACAGCGGATGATTAAGTTTGTGCTGCTTCCAGACTTTACAG 1327
 QY 1141 CAGATGCGTATTAACCGCGCTGTCAATTAATACGCGCGCGGAGAGGTGAACCTGACC 1200
 DB 1328 CAGATGCGTATTAACCGCGCTGTCAATTAATACGCGCGCGGAGAGGTGAACCTGACC 1387
 QY 1201 CTGCGAGATGTGAAGAGCGAAATGCGAGGCGCATGTGTGTTGAGGAGTTTACGCA 1260
 DB 1388 CTGCGAGATGTGAAGAGCGAAATGCGAGGCGCATGTGTGTTGAGGAGTTTACGCA 1447
 QY 1261 ATCGTAATGAAGACGCGATACCGCGCTGCAAGTTTGAAT 1300
 DB 1448 ATCGTAATGAAGACGCGATACCGCGCTGCAAGTTTGAAT 1487

RESULT 15
 ABK12514
 ID ABK12514 standard; DNA; 1299 BP.
 XX
 XX ABK12514;
 AC
 XX
 XX 05-JUN-2002 (first entry)
 DT
 XX
 XX DNA encoding phytase associated protein.
 DE
 XX
 XX KW Phytase; gene; da.
 XX

OS Unidentified.
 XX Key Location/Qualifiers
 FH CDS 1..1299
 FT /*tag= a
 FT /product= "Phytase associated protein"
 XX KR9086028-A.
 XX 15-DEC-1999.
 XX 25-MAY-1998; 98KR-00018810.
 XX 25-MAY-1998; 98KR-00018810.
 XX (WO0J-) WO0JIN CO LTD.
 XX Bae HD, Forceburgh CW, Goloben S, Cheng KJ;
 XX WPI: 2000-645078/62.
 XX P-PSDB; AAU77775.
 XX Novel phytase gene, recombinant phytase and usage thereof.
 XX Claim 1; Fig 2; 10pp; Korean.
 XX The invention relates to a novel phytase gene, a recombinant phytase gene
 CC and their uses. This sequence encodes a phytase associated protein,
 CC described in the invention
 XX
 SQ Sequence 1299 BP; 317 A; 345 C; 355 G; 282 T; 0 U; 0 Other;
 Query Match 97.4%; Score 1287; DB 3; Length 1299;
 Beg Local Similarity 99.5%; Pred. No. 0;
 Matches 1290; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 ATGAAAGGATCTTATCCCATTTTATCTCTTGTGATTCGTTAACCCCGCATCTGCA 60
 Db 1 ATGAAAGGATCTTATCCCATTTTATCTCTTGTGATTCGTTAACCCCGCATCTGCA 60
 QY 61 TTCGTCAGAGTGAAGCCGGAAGCTGAAAGTGTGATTTGTCAAGTGTGATGAT 120
 Db 61 TTCGTCAGAGTGAAGCCGGAAGCTGAAAGTGTGATTTGTCAAGTGTGATGAT 120
 QY 121 GTGCGTGTCCCAACCAAGGCCACCACTGATGAGATGTCACCCCGAAGCGATGGCA 180
 Db 121 GTGCGTGTCCCAACCAAGGCCACCACTGATGAGATGTCACCCCGAAGCGATGGCA 180
 QY 181 ACTGAGCGGTAACCTGGTTGGCTGACACCGCGAGTGGTGAAGCTATCGCTATCTC 240
 Db 181 ACTGAGCGGTAACCTGGTTGGCTGACACCGCGAGTGGTGAAGCTATCGCTATCTC 240
 QY 241 GGAATTACCAACCGCAGCTGTGAGCCGACGATTGCTGAGCAAAAGGGCTGCGCG 300
 Db 241 GGAATTACCAACCGCAGCTGTGAGCCGACGATTGCTGAGCAAAAGGGCTGCGCG 300
 QY 301 CAGTCTGTGATGAGTGGCAATTAATGTCGATGTCGAGCGATACCCGTTAAACAGCGCAA 360
 Db 301 CAGTCTGTGATGAGTGGCAATTAATGTCGATGTCGAGCGATACCCGTTAAACAGCGCAA 360
 QY 361 GCCTTGCCTGCGCGGCTGGCACTGACTGTGCAATTAACCGTATCCAGGAGATAG 420
 Db 361 GCCTTGCCTGCGCGGCTGGCACTGACTGTGCAATTAACCGTATCCAGGAGATAG 420
 QY 421 TCAGTCCCGATCGGTATTATTTATCTCTTAAACCTGGCGTTTGCCTGATTAACGCG 480
 Db 421 TCAGTCCCGATCGGTATTATTTATCTCTTAAACCTGGCGTTTGCCTGATTAACGCG 480
 QY 481 AACGTAAGTGAAGCGATCTGAGAGGCGAGAGGATCAATTGCTGATTAACCGGCGAT 540
 Db 481 AACGTAAGTGAAGCGATCTGAGAGGCGAGAGGATCAATTGCTGATTAACCGGCGAT 540
 QY 541 CGGCAAAAGGCGGTTTGGCAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600

Db 541 CGGCAAAAGGCGGTTTGGCAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600
 QY 601 CTTAAAGTGAAGAAACAGAGCAAAAGCTTTTCAATTAACGAGGATTAACCATCGGAATC 660
 Db 601 CTTAAAGTGAAGAAACAGAGCAAAAGCTTTTCAATTAACGAGGATTAACCATCGGAATC 660
 QY 661 AAGGTAGCGCGGCAATGCTCTAATTAACCGGTGGGTAAGCCCTGCAATGATGATGACG 720
 Db 661 AAGGTAGCGCGGCAATGCTCTAATTAACCGGTGGGTAAGCCCTGCAATGATGATGACG 720
 QY 721 GAGATATTTCTCTGCAACAAGCAAGGAAATGCGAGCGCGGGGTGGGAAGGATCAC 780
 Db 721 GAGATATTTCTCTGCAACAAGCAAGGAAATGCGAGCGCGGGGTGGGAAGGATCAC 780
 QY 781 GATTCAACACAGTGAACACCTTGTCTAAGTTTGTGATTAACGCGCAATTTTATTTCTACAA 840
 Db 781 GATTCAACACAGTGAACACCTTGTCTAAGTTTGTGATTAACGCGCAATTTTATTTCTACAA 840
 QY 841 CGCAGCGCAAGGTTTGGCCGCGAGCGCGCAACCCGTTAATTGATTTGATCATGGCAGCG 900
 Db 841 CGCAGCGCAAGGTTTGGCCGCGAGCGCGCAACCCGTTAATTGATTTGATCATGGCAGCG 900
 QY 901 TTGACGCCCATTCACCGCAAAAACAGCGCTGATGATGATTAACCATTCAGTCTG 960
 Db 901 TTGACGCCCATTCACCGCAAAAACAGCGCTGATGATGATTAACCATTCAGTCTG 960
 QY 961 TTTATTTCCGGAACAGTATCTGATGCAAAATCCGCGCGCGGACGTGAGCTCAACTG 1020
 Db 961 TTTATTTCCGGAACAGTATCTGATGCAAAATCCGCGCGCGGACGTGAGCTCAACTG 1020
 QY 1021 ACGCTCCCGGTGACCGGATTAACAGCCCGCAGGTGATGATGATGATGATGATGATGATGAT 1080
 Db 1021 ACGCTCCCGGTGACCGGATTAACAGCCCGCAGGTGATGATGATGATGATGATGATGATGAT 1080
 QY 1081 CGTGGGCTAAGCGATTAACAGCGAGTGAATTCAGGTTTCTGCTGATCTTCCAGACTTTACAG 1140
 Db 1081 CGTGGGCTAAGCGATTAACAGCGAGTGAATTCAGGTTTCTGCTGATCTTCCAGACTTTACAG 1140
 QY 1141 CAGATGGTGTATTAACAGCGCGCTGTCTAATTAATGCGCGCGGAGAGGTGAACCTGACC 1200
 Db 1141 CAGATGGTGTATTAACAGCGCGCTGTCTAATTAATGCGCGCGGAGAGGTGAACCTGACC 1200
 QY 1201 CTGGCAGGATGTAAGAGCGAAATGCGCAGGCGATGTTGTTGTTGAGGATTTTACGCAA 1260
 Db 1201 CTGGCAGGATGTAAGAGCGAAATGCGCAGGCGATGTTGTTGTTGAGGATTTTACGCAA 1260
 QY 1261 ATCGTGAATGAAGCAGCATACCGGCGTGCAGTTTG 1296
 Db 1261 ATCGTGAATGAAGCAGCATACCGGCGTGCAGTTTG 1296

Search completed: June 14, 2006, 10:31:32
 Job time : 903.103 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2006, 10:00:55 ; Search time 7399.95 Seconds
(without alignments)
11432.839 Million cell updates/sec

Title: US-10-601-319-1

Perfect score: 1322

Sequence: 1 atgaagaagcattcttaacc.....atcaccatcaccatcataa 1323

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_env:*
2: gb_pat:*
3: gb_ph:*
4: gb_pl:*
5: gb_pr:*
6: gb_ro:*
7: gb_sts:*
8: gb_sy:*
9: gb_un:*
10: gb_vl:*
11: gb_ov:*
12: gb_hcg:*
13: gb_in:*
14: gb_cm:*
15: gb_ba:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1332	100.0	1323	AR108133	AR108133 Sequence
2	1332	100.0	1323	AR127818	AR127818 Sequence
3	1332	100.0	1323	AR130956	AR130956 Sequence
4	1332	100.0	1323	AR492966	AR492966 Sequence
5	1332	100.0	1323	AR636185	AR636185 Sequence
6	1332	100.0	1323	AX052713	AX052713 Sequence
7	1332	100.0	1323	AX389866	AX389866 Sequence
8	1332	100.0	1323	AX356572	AX356572 Sequence
9	1287.8	97.4	1901	AR636190	AR636190 Sequence
10	1287.8	97.4	1901	AX356572	AX356572 Sequence
11	1287.8	97.4	1901	ECOAPPA	MS8708 Escherichia
12	1287.8	97.4	110000	U00096.10	Continuation (11 o
13	1287.8	97.4	110000	AP009048_10	Continuation (11 o
14	1286.2	97.3	1901	AR636188	AR636188 Sequence
15	1286.2	97.3	1901	AR636191	AR636191 Sequence
16	1286.2	97.3	1901	AX356570	AX356570 Sequence
17	1286.2	97.3	1901	AX356574	AX356574 Sequence
18	1286.2	97.3	1901	ECOACPHB	L03371 Escherichia

19	1286.2	97.3	110000	15	CP000038_10	Continuation (11 o
20	1284.6	97.2	1315	15	AF537219	AF537219 Escherich
21	1284.6	97.2	1901	15	ECOACPHD	L03372 Escherichia
22	1284.6	97.2	1901	15	ECOACPHD	L03373 Escherichia
23	1284.6	97.2	1901	15	ECOACPHD	L03374 Escherichia
24	1284.6	97.2	1901	15	ECOACPHD	L03375 Escherichia
25	1283	97.0	1901	2	AR636189	AR636189 Sequence
26	1283	97.0	1901	2	AX356571	AX356571 Sequence
27	1283	97.0	1901	15	ECOACPHA	L03370 Escherichia
28	1283	97.0	110000	15	CP000036_22	Continuation (23 o
29	1282.2	97.0	3470	2	AX042376	AX042376 Sequence
30	1282.2	97.0	4060	2	AX042374	AX042374 Sequence
31	1282.2	97.0	6116	2	AX042375	AX042375 Sequence
32	1282.2	97.0	6708	2	AX042373	AX042373 Sequence
33	1282.2	97.0	17732	2	AX042378	AX042378 Sequence
34	1282.2	97.0	20623	2	AX042372	AX042372 Sequence
35	1273.4	96.3	110000	15	AE005674_10	Continuation (11 o
36	1271.8	96.2	110000	15	AE014073_10	Continuation (11 o
37	1270.2	96.1	5421	2	AX042377	AX042377 Sequence
38	1263.8	95.6	1489	2	AR277753	AR277753 Sequence
39	1263.8	95.6	1489	15	AY96073	AY96073 Escherich
40	1263	95.5	1486	2	AR630587	AR630587 Sequence
41	1251.8	94.7	1486	2	AR630588	AR630588 Sequence
42	1246.6	94.3	110000	15	AE005174_12	Continuation (13 o
43	1246.6	94.3	110000	15	AE005174_13	Continuation (14 o
44	1246.6	94.3	110000	15	BA000007_12	Continuation (13 o
45	1246	94.3	110000	15	AE014075_10	Continuation (11 o

ALIGNMENTS

RESULT 1	AR108133	1323 bp	DNA	linear	PAT 14-FEB-2001
LOCUS	AR108133				
DEFINITION	Sequence 1 from patent US 6110719.				
ACCESSION	AR108133				
VERSION	AR108133.1 GI:12823620				
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 1323)				
TITLE	Kretz, K.				
JOURNAL	Phyrase				
FEATURES	Patent: US 6110719-A 1 29-AUG-2000;				
source	Location/Qualifiers				
	1..1323				
	/organism="unknown"				
	/mol_type="unassigned DNA"				
ORIGIN					
Query Match	100.0%; Score 1322; DB 2; Length 1323;				
Best Local Similarity	100.0%; Pred. No. 0;				
Matches 1323; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
QY	1 ATGAAGGATTTATCCATTTTATCTTCTGATTCGTTAACCCGGAATCTGCA	60			
DB	1 ATGAAGGATTTATCCATTTTATCTTCTGATTCGTTAACCCGGAATCTGCA	60			
QY	61 TTGCGTCAGATGAGCCGAGCTGAAAGTGTGGATGATGATGATGATGAT	120			
DB	61 TTGCGTCAGATGAGCCGAGCTGAAAGTGTGGATGATGATGATGATGAT	120			
QY	121 GTGCGTCGTCACCAAGGACGCACTGATGATGATGATGATGATGATGAT	180			
DB	121 GTGCGTCGTCACCAAGGACGCACTGATGATGATGATGATGATGATGAT	180			
QY	181 ACTGCGCGGTTAACTGCGTTGCTGACACCGCGGTTGAGTGAATGCTAT	240			
DB	181 ACTGCGCGGTTAACTGCGTTGCTGACACCGCGGTTGAGTGAATGCTAT	240			
QY	241 GGACATTACCAAGCAGCGCTCTGTAGCCAGATTCGTGGCAAAAAGGCG	300			

Db 241 GGAACATTACCAACGACGAGTCTGGTAGCCAGCGAATTCCTGCGCAAAAAGGCGTCCCG 300
QY 301 CAGTCTGGTCAAGTGGCGGATTATTTGGTAGTTCGACGAGCGTAAACCGGTAACAGCCGAA 360
Db 301 CAGTCTGGTCAAGTGGCGGATTATTTGGTAGTTCGACGAGCGTAAACCGGTAACAGCCGAA 360
QY 361 GCGTTCGCGCGGCGGTGGCACTGATCTGTGCAATTAACCGTACATACCAGGCAATACG 420
Db 361 GCGTTCGCGCGGCGGTGGCACTGATCTGTGCAATTAACCGTACATACCAGGCAATACG 420
QY 421 TCCAGTCCCGATCCGTTATTTAATCTCTAAAAAACCTGCGTTCGCAATGATTAACGCG 480
Db 421 TCCAGTCCCGATCCGTTATTTAATCTCTAAAAAACCTGCGTTCGCAATGATTAACGCG 480
QY 481 AACGTGACTGACGCGATCTCTGACGAGGCGAGAGGGTCAATTGCTGACTTTACCGGCGAT 540
Db 481 AACGTGACTGACGCGATCTCTGACGAGGCGAGAGGGTCAATTGCTGACTTTACCGGCGAT 540
QY 541 CGGCAAAAGGCGTTTCGCGAATGGAAGGGGTCTTAATTTCCGCAATCAAACTTGTGC 600
Db 541 CGGCAAAAGGCGTTTCGCGAATGGAAGGGGTCTTAATTTCCGCAATCAAACTTGTGC 600
QY 601 CTTAAACGTGAGAAAACAGGACGAAAGCTGTTCAATTAACGACGCAATTAACATCGAACTC 660
Db 601 CTTAAACGTGAGAAAACAGGACGAAAGCTGTTCAATTAACGACGCAATTAACATCGAACTC 660
QY 661 AAGGTGAGCGCGGACCAATATCTCAATTAACCGGTGGTAAAGCTCGCATATCTGACG 720
Db 661 AAGGTGAGCGCGGACCAATATCTCAATTAACCGGTGGTAAAGCTCGCATATCTGACG 720
QY 721 GAGATATTTCTCTCTCAACAAGCAGAGGAAATGCGAGCCGAGGTTGGGAAAGATCAC 780
Db 721 GAGATATTTCTCTCTCAACAAGCAGAGGAAATGCGAGCCGAGGTTGGGAAAGATCAC 780
QY 781 GATTCAACACGATGGAACACTTGTCTAAGTTTGCAATACGCGCAATTTATTTGCTACAA 840
Db 781 GATTCAACACGATGGAACACTTGTCTAAGTTTGCAATACGCGCAATTTATTTGCTACAA 840
QY 841 CGCAAGCCGAGAGTGGCCGCGAGCCGCGCAACCCCGTTATTTGAAATTTGATGCGACG 900
Db 841 CGCAAGCCGAGAGTGGCCGCGAGCCGCGCAACCCCGTTATTTGAAATTTGATGCGACG 900
QY 901 TTGAGCGCCCATTCACCGCAAAAACAGGCGTATGATGATCAATTAACCACTTCAGTACG 960
Db 901 TTGAGCGCCCATTCACCGCAAAAACAGGCGTATGATGATCAATTAACCACTTCAGTACG 960
QY 961 TTTATTTGCGGACACGATTAATACTGAGCAATCTCGCGGCGCACTGAGGCTCAACTCG 1020
Db 961 TTTATTTGCGGACACGATTAATACTGAGCAATCTCGCGGCGCACTGAGGCTCAACTCG 1020
QY 1021 AAGCTTCCCGGTCAAGCCGGAATAACGCGCGCAAGTGTGAATCTGTGTTGAAGCTCG 1080
Db 1021 AAGCTTCCCGGTCAAGCCGGAATAACGCGCGCAAGTGTGAATCTGTGTTGAAGCTCG 1080
QY 1081 CGTCCGCTAAGGATTAACGACGATGATTCAGTTCCTGAGCTTCAAGCTTCAAG 1140
Db 1081 CGTCCGCTAAGGATTAACGACGATGATTCAGTTCCTGAGCTTCAAGCTTCAAG 1140
QY 1141 CAGATGCGTATTAACGCGCTGTCTAATTAATACGCGCGCGGAGAGGTGAACCTGAC 1200
Db 1141 CAGATGCGTATTAACGCGCTGTCTAATTAATACGCGCGCGGAGAGGTGAACCTGAC 1200
QY 1201 CCGGACGATGTAAGAGAGGAAATGCGAGGCGCAATGTTGTTGCGAGGTTTACGCA 1260
Db 1201 CCGGACGATGTAAGAGAGGAAATGCGAGGCGCAATGTTGTTGCGAGGTTTACGCA 1260
QY 1261 ATCGTAATGAAGACGACATACCGCGGTCAAGTTGAGATCTCATCAACATCAATCAC 1320
Db 1261 ATCGTAATGAAGACGACATACCGCGGTCAAGTTGAGATCTCATCAACATCAATCAC 1320
QY 1321 TAA 1323
Db 1321 TAA 1323

Db 1321 TAA 1323
RESULT 2
LOCUS AR127818 1323 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6183740.
ACCESSION AR127818
VERSION AR127818.1 GI:14115480
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1323)
AUTHORS Short, J.M. and Kreitz, K.A.
TITLE Recombinant bacterial phycases and uses thereof
JOURNAL Patent: US 6183740-A 1 06-FEB-2001;
FEATURES
source location/Qualifiers
1..1323
/organism="Unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 1322; DB 2; Length 1323;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAAGGATCTTAATCCCAATTTTATCTCTGATTCGGTAAACCCGCAATCTGCA 60
Db 1 ATGAAAGGATCTTAATCCCAATTTTATCTCTGATTCGGTAAACCCGCAATCTGCA 60
QY 61 TTCGCTCAGATGAGCCGGAAGCTGGAAGTGTGATGATGATGATGATGATGATGATGAT 120
Db 61 TTCGCTCAGATGAGCCGGAAGCTGGAAGTGTGATGATGATGATGATGATGATGATGAT 120
QY 121 GTGCGTCTCAACCAAGCCGCAAGCACTGATGAGATGATGATGATGATGATGATGATGAT 180
Db 121 GTGCGTCTCAACCAAGCCGCAAGCACTGATGAGATGATGATGATGATGATGATGATGAT 180
QY 181 AACTGCGCGGTAAACCTGAGTGGTGAACACCGCGGATGATGATGATGATGATGATGATGAT 240
Db 181 AACTGCGCGGTAAACCTGAGTGGTGAACACCGCGGATGATGATGATGATGATGATGATGAT 240
QY 241 GGAACATTAACAGCGCAAGCTCTGGTACCGACGATGCTGCGCAAAAAGGCGTCCG 300
Db 241 GGAACATTAACAGCGCAAGCTCTGGTACCGACGATGCTGCGCAAAAAGGCGTCCG 300
QY 301 CAGTCTGATGAGTGGCGATTAATGCTGATGATGATGATGATGATGATGATGATGATGAT 360
Db 301 CAGTCTGATGAGTGGCGATTAATGCTGATGATGATGATGATGATGATGATGATGATGAT 360
QY 361 GCGTTCGCGCGGCGGTGGCACTGATCTGTGCAATTAACCGTACATACCAGGCAATACG 420
Db 361 GCGTTCGCGCGGCGGTGGCACTGATCTGTGCAATTAACCGTACATACCAGGCAATACG 420
QY 421 TCCAGTCCCGATCCGTTATTTAATCTCTAAAAAACCTGCGTTCGCAATGATTAACGCG 480
Db 421 TCCAGTCCCGATCCGTTATTTAATCTCTAAAAAACCTGCGTTCGCAATGATTAACGCG 480
QY 481 AACGTGACTGACGCGATCTCTGACGAGGCGAGAGGGTCAATTGCTGACTTTACCGGCGAT 540
Db 481 AACGTGACTGACGCGATCTCTGACGAGGCGAGAGGGTCAATTGCTGACTTTACCGGCGAT 540
QY 541 CGGCAAAAGGCGTTTCGCGAATGGAAGGGGTCTTAATTTCCGCAATCAAACTTGTGC 600
Db 541 CGGCAAAAGGCGTTTCGCGAATGGAAGGGGTCTTAATTTCCGCAATCAAACTTGTGC 600
QY 601 CTTAAACGTGAGAAAACAGGACGAAAGCTGTTCAATTAACGACGCAATTAACATCGAACTC 660
Db 601 CTTAAACGTGAGAAAACAGGACGAAAGCTGTTCAATTAACGACGCAATTAACATCGAACTC 660
QY 661 AAGGTGAGCGCGGACCAATATCTCAATTAACCGGTGGTAAAGCTCGCATATCTGACG 720
Db 661 AAGGTGAGCGCGGACCAATATCTCAATTAACCGGTGGTAAAGCTCGCATATCTGACG 720

Db 661 AAGTAGAGCCGACATATCTCATTAACCGGTGCGGTAAAGCTCGATCAATGCTGAG 720
Qy 721 GAGATATTTCTCTGCAACAAGCAAGGAAATCCGAGCCGGGTGGGGAAGATCACC 780
Db 721 GAGATATTTCTCTGCAACAAGCAAGGAAATCCGAGCCGGGTGGGGAAGATCACC 780
Qy 781 GATTACACACAGTGAACAACCTTGCTAAGTTTGATTAAGCGCAATTTTATTTGCTACAA 840
Db 781 GATTACACACAGTGAACAACCTTGCTAAGTTTGATTAAGCGCAATTTTATTTGCTACAA 840
Qy 841 CGCAGCCGACAGAGTTGTCGCGACGCGCCACCCCGTTATGTGATTCATGAGAGCG 900
Db 841 CGCAGCCGACAGAGTTGTCGCGACGCGCCACCCCGTTATGTGATTCATGAGAGCG 900
Qy 901 TTGACGCCCCCTATCACCAGCAAAAACAGCGGTATGTGTGACATTAACCACTCACTG 960
Db 901 TTGACGCCCCCTATCACCAGCAAAAACAGCGGTATGTGTGACATTAACCACTCACTG 960
Qy 961 TTTATTTGCGGACACGATCTAATCTGGCAAACTGGGCGGCGACATGAGGCTCACTGG 1020
Db 961 TTTATTTGCGGACACGATCTAATCTGGCAAACTGGGCGGCGACATGAGGCTCACTGG 1020
Qy 1021 ACGCTTCCCGGTACGCGGATTAACAAGCCGACAGTGTGAATGTGTGTTGAACGCTGG 1080
Db 1021 ACGCTTCCCGGTACGCGGATTAACAAGCCGACAGTGTGAATGTGTGTTGAACGCTGG 1080
Qy 1081 CGTGGCTTAAGGATTAACAGCCAGTGTGATTCAGGTTTCGCTGGTCTTCCAGCTTTACAG 1140
Db 1081 CGTGGCTTAAGGATTAACAGCCAGTGTGATTCAGGTTTCGCTGGTCTTCCAGCTTTACAG 1140
Qy 1141 CAGATCGGATTAACAGCCGCTGTCTAATTAATACGCCCGCGGAAAGGTGAACCTGACC 1200
Db 1141 CAGATCGGATTAACAGCCGCTGTCTAATTAATACGCCCGCGGAAAGGTGAACCTGACC 1200
Qy 1201 CTGGCAGATGTGAAGAGCGAAATGCGAGGCGATGTGTGCTTGGCAGGTTTACGCAA 1260
Db 1201 CTGGCAGATGTGAAGAGCGAAATGCGAGGCGATGTGTGCTTGGCAGGTTTACGCAA 1260
Qy 1261 ATCGTGAATGAAGACGCAATACCGGCGTGCAGTTGATTCATCAACATCAACATCAC 1320
Db 1261 ATCGTGAATGAAGACGCAATACCGGCGTGCAGTTGATTCATCAACATCAACATCAC 1320
Qy 1321 TAA 1323
Db 1321 TAA 1323

RESULT 3
ARI30956
LOCUS ARI30956 1323 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6190897.
ACCESSION ARI30956
VERSION ARI30956.1 GI:14119281
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1323)
AUTHORS Kretz, K.
TITLE Phycase
JOURNAL Patent: US 6190897-A 1 20-PBB-2001;
FEATURES
Location/Qualifiers
1..1323
/organism="unknown"
/mol_type="unassigned DNA"

Query Match 100.0%; Score 1322; DB 2; Length 1323;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAGGAGATCTTAATCCCATTTTATCTTGTGATTCGTTAAGCCGGAATCTGCA 60
|||||
Db 1 ATGAAGGAGATCTTAATCCCATTTTATCTTGTGATTCGTTAAGCCGGAATCTGCA 60

Db 1 ATGAAGGAGATCTTAATCCCATTTTATCTTGTGATTCGTTAAGCCGGAATCTGCA 60
Qy 61 TTGCTCAGAGTGAAGCCGAGCTGAAGCTGAAAGTGTGTGATTTGTCAGTCTAGGT 120
Db 61 TTGCTCAGAGTGAAGCCGAGCTGAAGCTGAAAGTGTGTGATTTGTCAGTCTAGGT 120
Qy 121 GTGGCTGCTCCCAACCAAGGCCACGCAACTGATGCAAGATGTCACCCCCAACGATGGCA 180
Db 121 GTGGCTGCTCCCAACCAAGGCCACGCAACTGATGCAAGATGTCACCCCCAACGATGGCA 180
Qy 181 ACCGTGCGGGTAAACTGGGTTGGCTGACCGCGGNGTGTGAGCTAATCGCTATCTC 240
Db 181 ACCGTGCGGGTAAACTGGGTTGGCTGACCGCGGNGTGTGAGCTAATCGCTATCTC 240
Qy 241 GGAATTAACCAACGCCAGGCTGTGATGCGACGGAATTTGCGGCAAAAAGGCTGCGCG 300
Db 241 GGAATTAACCAACGCCAGGCTGTGATGCGACGGAATTTGCGGCAAAAAGGCTGCGCG 300
Qy 301 CAGTCTGGTCAAGTGGCGAATTAATGCTGAATGTGACAGACGTTACCCGTAAACAGGCGAA 360
Db 301 CAGTCTGGTCAAGTGGCGAATTAATGCTGAATGTGACAGACGTTACCCGTAAACAGGCGAA 360
Qy 361 GCCTTCGCGCGGCTGGCACTGACAGTGTGAATTAACGTAATTAACCCAGGCAATAG 420
Db 361 GCCTTCGCGCGGCTGGCACTGACAGTGTGAATTAACGTAATTAACCCAGGCAATAG 420
Qy 421 TCAGTCCCGATCCGTTATTTAATCTCTAATAAATGCGGTTTGCCAACTGATTAACGCG 480
Db 421 TCAGTCCCGATCCGTTATTTAATCTCTAATAAATGCGGTTTGCCAACTGATTAACGCG 480
Qy 481 AACGTGATGAGCGCATCTCAGCAGAGGCGAGAGGTCATTTGCTGATTCACCGGCAAT 540
Db 481 AACGTGATGAGCGCATCTCAGCAGAGGCGAGAGGTCATTTGCTGATTCACCGGCAAT 540
Qy 541 CGGCAAAACGCGCTTTCGCGAACTGGAACGGGTGCTTAATTTTCGCAATCAAACTTGTGC 600
Db 541 CGGCAAAACGCGCTTTCGCGAACTGGAACGGGTGCTTAATTTTCGCAATCAAACTTGTGC 600
Qy 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGACAGGATTAACATCGGAATC 660
Db 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGACAGGATTAACATCGGAATC 660
Qy 661 AAGGTGAGCGCGCAATGTCTCATTAACCGGTGGGTAAAGCTGTCATATGCTGAGC 720
Db 661 AAGGTGAGCGCGCAATGTCTCATTAACCGGTGGGTAAAGCTGTCATATGCTGAGC 720
Qy 721 GAGATATTTCTCTGCAACAAGCAAGGAAATGCGAGCGGGGTGGGGAAGATCACC 780
Db 721 GAGATATTTCTCTGCAACAAGCAAGGAAATGCGAGCGGGGTGGGGAAGATCACC 780
Qy 781 GATTACACACAGTGAACAACCTTGCTAAGTTTGATTAACGCGCAATTTTATTTGCTACAA 840
Db 781 GATTACACACAGTGAACAACCTTGCTAAGTTTGATTAACGCGCAATTTTATTTGCTACAA 840
Qy 841 CGCAGCCGACAGAGTTGTCGCGACGCGCCACCCCGTTATGTGATTCATGAGAGCG 900
Db 841 CGCAGCCGACAGAGTTGTCGCGACGCGCCACCCCGTTATGTGATTCATGAGAGCG 900
Qy 901 TTGACGCCCCCTATCACCAGCAAAAACAGCGGTATGTGTGACATTAACCACTCACTG 960
Db 901 TTGACGCCCCCTATCACCAGCAAAAACAGCGGTATGTGTGACATTAACCACTCACTG 960
Qy 961 TTTATTTGCGGACACGATCTAATCTGGCAAACTGGGCGGCGACATGAGGCTCACTGG 1020
Db 961 TTTATTTGCGGACACGATCTAATCTGGCAAACTGGGCGGCGACATGAGGCTCACTGG 1020
Qy 1021 ACGCTTCCCGGTACGCGGATTAACAAGCCGACAGTGTGAATGTGTGTTGAACGCTGG 1080
Db 1021 ACGCTTCCCGGTACGCGGATTAACAAGCCGACAGTGTGAATGTGTGTTGAACGCTGG 1080
Qy 1081 CGTGGCTTAAGGATTAACAGCCAGTGTGATTCAGGTTTCGCTGGTCTTCCAGCTTTACAG 1140
Db 1081 CGTGGCTTAAGGATTAACAGCCAGTGTGATTCAGGTTTCGCTGGTCTTCCAGCTTTACAG 1140

QY 1141 CAGATGCGTGAATAAAGCGCGCTGTCATTAAATAGCCGCCCGGAGAGGTGAAGTGAACC 1200
DB 1141 CAGATGCGTGAATAAAGCGCGCTGTCATTAAATAGCCGCCCGGAGAGGTGAAGTGAACC 1200
QY 1201 CTGGCAGGATGTGAAGAGCGAATGCGAGGCGATGTTGTTGCTGGCAGGTTTACGCA 1260
DB 1201 CTGGCAGGATGTGAAGAGCGAATGCGAGGCGATGTTGTTGCTGGCAGGTTTACGCA 1260
QY 1261 ATCGTAATGAAGACGATACCGGCGGTGCAATTGAGATCTCATCACCATCATCAC 1320
DB 1261 ATCGTAATGAAGACGATACCGGCGGTGCAATTGAGATCTCATCACCATCATCAC 1320
QY 1321 TAA 1323
DB 1321 TAA 1323

RESULT 4
AR492966
LOCUS AR492966 1323 bp DNA linear PAT 15-MAY-2004
DEFINITION Sequence 1 from patent US 6720014.
ACCESSION AR492966
VERSION AR492966.1 GI:47264255
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1323)
AUTHORS Short, J.M. and Kretz, K.A.
TITLE Phylase- containing foodstuffs and methods of making and using them
JOURNAL Patent: US 6720014-A 1 13-APR-2004;
Diversa Corporation; San Diego, CA

FEATURES
source
1..1323
/organism="Unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 1322; DB 2; Length 1323;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAGGATCTTAATCCATTTTATCTCTTGATTCGGTAAACCGCAATCTGCA 60
DB 1 ATGAAAGGATCTTAATCCATTTTATCTCTTGATTCGGTAAACCGCAATCTGCA 60
QY 61 TTCGTCAGAGTGAGCCCGAGCTGAAGCTGAAGAGTGTGATTTGATGTCATGAT 120
DB 61 TTCGTCAGAGTGAGCCCGAGCTGAAGAGTGTGATTTGATGTCATGATGAT 120
QY 121 GTGCGTGTCTCCACCAAGCCGACGCAATGATGTCAGATGTCAACCCGACGATGGCA 180
DB 121 GTGCGTGTCTCCACCAAGCCGACGCAATGATGTCAGATGTCAACCCGACGATGGCA 180
QY 181 ACCGCGCGGTAACCTGGGTTGCTGACACCGGAGGTGTGAGTATGCGCTATCTC 240
DB 181 ACCGCGCGGTAACCTGGGTTGCTGACACCGGAGGTGTGAGTATGCGCTATCTC 240
QY 241 GGACATTAACCAACGCGCAGCGTCTGTGAGCCGACGATTTGCTGGCAAAAAGGCTGCCG 300
DB 241 GGACATTAACCAACGCGCAGCGTCTGTGAGCCGACGATTTGCTGGCAAAAAGGCTGCCG 300
QY 301 CAGTTCGTGTCAGGTGCGGATTTTGTGATGTGACAGAGCTTACCCGTAACACGCGCA 360
DB 301 CAGTTCGTGTCAGGTGCGGATTTTGTGATGTGACAGAGCTTACCCGTAACACGCGCA 360
QY 361 GCCTTGCGCGCGGCTGGCAGCTGACGTGTGCAATTAACGTAACACGAGGAGATAG 420
DB 361 GCCTTGCGCGCGGCTGGCAGCTGACGTGTGCAATTAACGTAACACGAGGAGATAG 420
QY 421 TTCAGTCCCGATCGGTATTATCTCTTAAACCTGGCGTTTGCACATGATTAACGCG 480

DB 421 TCAGTCCCGATCGGTATTATCTCTTAAACCTGGCGTTTGCACATGATTAACGCG 480
QY 481 AACGTGACTGACGGGATCTTCAAGAGGCGAGAGGTCATTTGTGACTTTACCGGGAT 540
DB 481 AACGTGACTGACGGGATCTTCAAGAGGCGAGAGGTCATTTGTGACTTTACCGGGAT 540
QY 541 CGGCAACCGGCTTTCGCAACTGGAACGGGTGCTTAATTTTCCGATCAACCTTGTGC 600
DB 541 CGGCAACCGGCTTTCGCAACTGGAACGGGTGCTTAATTTTCCGATCAACCTTGTGC 600
QY 601 CTTAAACGTGAAGAACAGAGCAAGCTTTCATTAAACGAGGATTAACATTCGAACTC 660
DB 601 CTTAAACGTGAAGAACAGAGCAAGCTTTCATTAAACGAGGATTAACATTCGAACTC 660
QY 661 AAGGTGAGCGCGGACATGTCATTAAACCGGTGCGGTGACCTTCGATCAATGTCAGC 720
DB 661 AAGGTGAGCGCGGACATGTCATTAAACCGGTGCGGTGACCTTCGATCAATGTCAGC 720
QY 721 GAGATTAATTTCTCTGCAACAGCAAGGAAATGCGGAGCCGGGGTGGGAAAGATCAC 780
DB 721 GAGATTAATTTCTCTGCAACAGCAAGGAAATGCGGAGCCGGGGTGGGAAAGATCAC 780
QY 781 GATTCAACCAAGTGAACCACTTGTCTAAGTTTGATTAACGCGCAATTTTATTTGCTACAA 840
DB 781 GATTCAACCAAGTGAACCACTTGTCTAAGTTTGATTAACGCGCAATTTTATTTGCTACAA 840
QY 841 CGCAGCGCAAGAGTTTGCCTCGGAGCCGCGCCACCCCGTTATTTGATTTGATTAACGCGC 900
DB 841 CGCAGCGCAAGAGTTTGCCTCGGAGCCGCGCCACCCCGTTATTTGATTTGATTAACGCGC 900
QY 901 TTGACGCGCCCATTCACGCGCAAAAACAGGCGATGATGATGATTAACCATTCATGATCTG 960
DB 901 TTGACGCGCCCATTCACGCGCAAAAACAGGCGATGATGATGATTAACCATTCATGATCTG 960
QY 961 TTATATGCGCGGACAGATTAATCTGCGCAAAATCTCGCGCGCACTGAGCTCAACTGG 1020
DB 961 TTATATGCGCGGACAGATTAATCTGCGCAAAATCTCGCGCGCACTGAGCTCAACTGG 1020
QY 1021 ACGCTTCCCGGTTCAGCCGGATTAACAGCCCGCAGGTGTGAACCTGTTGAACGCTGG 1080
DB 1021 ACGCTTCCCGGTTCAGCCGGATTAACAGCCCGCAGGTGTGAACCTGTTGAACGCTGG 1080
QY 1081 CGTGGCTTAACGATTAACAGCAGTGTGATTCAGGTTTCTGCTGCTTCCAGACTTTACAG 1140
DB 1081 CGTGGCTTAACGATTAACAGCAGTGTGATTCAGGTTTCTGCTGCTTCCAGACTTTACAG 1140
QY 1141 CAGATGCGTGAATAAAGCGCGCTGTCTTAATAATAGCCGCCCGGAGAGGTGAACCTGACC 1200
DB 1141 CAGATGCGTGAATAAAGCGCGCTGTCTTAATAATAGCCGCCCGGAGAGGTGAACCTGACC 1200
QY 1201 CTGGCAGGATGTGAAGAGCGAATGCGAGGCGATGTTGCTGGCAGGTTTACGCA 1260
DB 1201 CTGGCAGGATGTGAAGAGCGAATGCGAGGCGATGTTGCTGGCAGGTTTACGCA 1260
QY 1261 ATCGTAATGAAGACGATACCGGCGGTGCAATTGAGATCTCATCACCATCATCAC 1320
DB 1261 ATCGTAATGAAGACGATACCGGCGGTGCAATTGAGATCTCATCACCATCATCAC 1320
QY 1321 TAA 1323
DB 1321 TAA 1323

RESULT 5
AR636185
LOCUS AR636185 1323 bp DNA linear PAT 20-APR-2005
DEFINITION Sequence 1 from patent US 6853365.
ACCESSION AR636185
VERSION AR636185.1 GI:62768868
KEYWORDS
SOURCE
ORGANISM Unknown.
unclassified.

REFERENCE 1 (bases 1 to 1323)
 AUTHORS Short, J.M., Kretz, K.A., Gray, K.A., Barton, N.R., Garrett, J.B.,
 O'Donoghue, B. and Mathur, B.J.
 TITLE Recombinant bacterial phytases and uses thereof
 JOURNAL Patent: US 6853365-A 1 15-FEB-2005;
 Diversa Corporation; San Diego, CA
 FEATURES
 source 1..1323
 /organism="unknown"
 /mol_type="genomic DNA"
 ORIGIN
 Query Match 100.0%; Score 1322; DB 2; Length 1323;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAGCATCTTAAATCCATTTTATCTCTTCTGATTCCTGTAACCCCGCAATCTGCA 60
 DB 1 ATGAAAGCATCTTAAATCCATTTTATCTCTTCTGATTCCTGTAACCCCGCAATCTGCA 60
 QY 61 TTCGCTCAAGTGAAGCCGAGCTGAAAGCTGGAAGTGTGTGATGTCAGTCTCATGCT 120
 DB 61 TTCGCTCAAGTGAAGCCGAGCTGAAAGCTGGAAGTGTGTGATGTCAGTCTCATGCT 120
 QY 121 GTGCGTCTCAACCAAGGCGACGCAATGATGCAAGATGTCAACCCGACAGCATGGCCA 180
 DB 121 GTGCGTCTCAACCAAGGCGACGCAATGATGCAAGATGTCAACCCGACAGCATGGCCA 180
 QY 181 ACCTGGCCGGTAAACTGGGTTTGGCTGACACCGCGAGTGTGAGTAAATCCGCTATCTC 240
 DB 181 ACCTGGCCGGTAAACTGGGTTTGGCTGACACCGCGAGTGTGAGTAAATCCGCTATCTC 240
 QY 241 GGAACATTAACAACCCAGGCTGTGTAAGCCGACGATGCTGGCGAAAAAGGCTGCCG 300
 DB 241 GGAACATTAACAACCCAGGCTGTGTAAGCCGACGATGCTGGCGAAAAAGGCTGCCG 300
 QY 301 CAGTCTGATCAGATGCGCATTAATGCTGATGTCAGAGCGATACCCGTAACAGGCGAA 360
 DB 301 CAGTCTGATCAGATGCGCATTAATGCTGATGTCAGAGCGATACCCGTAACAGGCGAA 360
 QY 361 GCGTTTCGCGCGCGGCTGACCTGACCTGCAATTAACCGTACATACCCAGGCAATAG 420
 DB 361 GCGTTTCGCGCGCGGCTGACCTGACCTGCAATTAACCGTACATACCCAGGCAATAG 420
 QY 421 TCCAGTCCCGATCCGTTATTTAATCTCTTAATACTGCGGTTTCCCACTGATTAACGCG 480
 DB 421 TCCAGTCCCGATCCGTTATTTAATCTCTTAATACTGCGGTTTCCCACTGATTAACGCG 480
 QY 481 AACGTGATGACGCGATCTCTGACAGGAGGAGGATCAATGCTGACTTTACCGGGCAT 540
 DB 481 AACGTGATGACGCGATCTCTGACAGGAGGAGGATCAATGCTGACTTTACCGGGCAT 540
 QY 541 CGGCAACGCGCTTTCGCGAATGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600
 DB 541 CGGCAACGCGCTTTCGCGAATGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600
 QY 601 CTTAAACGTGAAAAAGGACGAAAGCTGTTCAATTAACGACGCAATTAACATCGAACTC 660
 DB 601 CTTAAACGTGAAAAAGGACGAAAGCTGTTCAATTAACGACGCAATTAACATCGAACTC 660
 QY 661 AAGGTGAGGCGGACAAATGTCATTAACGGTGGGTGAAGCTGCAATCAATGCTGACG 720
 DB 661 AAGGTGAGGCGGACAAATGTCATTAACGGTGGGTGAAGCTGCAATCAATGCTGACG 720
 QY 721 GAGATATTTCTCTGCAACAAAGCAAGGAAATGCGAGCCGAGGTGGGAAAGATCAC 780
 DB 721 GAGATATTTCTCTGCAACAAAGCAAGGAAATGCGAGCCGAGGTGGGAAAGATCAC 780
 QY 781 GATTCAACACAGTGAACACTCTTGCTAAGTTGCAATACGCGCAATTTATTTGCTACA 840
 DB 781 GATTCAACACAGTGAACACTCTTGCTAAGTTGCAATACGCGCAATTTATTTGCTACA 840
 QY 841 CGCAGCGCAGAGTTGCCGCGACGCGCGCACCCGTTATTGATTTGATGACGAGG 900

DB 841 CGCAGCGCAGAGTTGCCGCGACGCGCGCACCCGTTATTGATTTGATTCATGCGAGG 900
 QY 901 TTGAGCGCCCATCAACCGCAAAACAGCGGTATGTGTGACATTAACCACTTCAGTACTG 960
 DB 901 TTGAGCGCCCATCAACCGCAAAACAGCGGTATGTGTGACATTAACCACTTCAGTACTG 960
 QY 961 TTTTATTTGCGGACACGATCTAATCTGGCAATTTCTCGCGCGGCGCATGGAAGTCAACTG 1020
 DB 961 TTTTATTTGCGGACACGATCTAATCTGGCAATTTCTCGCGCGGCGCATGGAAGTCAACTG 1020
 QY 1021 ACGCTTCCCGGTCAAGCCGGAATTAACAACGCGCGACAGGAGTGAATCTGTTTGAAGCTG 1080
 DB 1021 ACGCTTCCCGGTCAAGCCGGAATTAACAACGCGCGACAGGAGTGAATCTGTTTGAAGCTG 1080
 QY 1081 CGTGGCTAAGCGATTAACAACGCGGTGATTCAGTTCCTGCTGCTTCCAGACTTACAG 1140
 DB 1081 CGTGGCTAAGCGATTAACAACGCGGTGATTCAGTTCCTGCTGCTTCCAGACTTACAG 1140
 QY 1141 CAGATGCGTGAATAAACCGCGCTGCTATTAATTAACGCGCGCGGAGAGGTGAACCTGACC 1200
 DB 1141 CAGATGCGTGAATAAACCGCGCTGCTATTAATTAACGCGCGCGGAGAGGTGAACCTGACC 1200
 QY 1201 CTGGCAGAGTGAAGAGGCAAAATGCGAGGCGATGTTGTTGGTGGCAGGTTTACGCA 1260
 DB 1201 CTGGCAGAGTGAAGAGGCAAAATGCGAGGCGATGTTGTTGGTGGCAGGTTTACGCA 1260
 QY 1261 ATCGTAATGAAGACGCAATACCGCGCTGCAATTTGATTCATCAACCATCACTCAC 1320
 DB 1261 ATCGTAATGAAGACGCAATACCGCGCTGCAATTTGATTCATCAACCATCACTCAC 1320
 QY 1321 TAA 1323
 DB 1321 TAA 1323

RESULT 5
 AX052713
 LOCUS 1323 bp DNA linear PART 12-JUN-2001
 DEFINITION Sequence 1 from Patent WO071728.
 ACCESSION AX052713
 VERSION
 KEYWORDS
 SOURCES
 ORGANISM Escherichia coli
 Escherichia coli
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 Enterobacteriaceae; Escherichia.
 REFERENCE 1
 AUTHORS Short, J.M. and Kretz, K.A.
 TITLE Recombinant bacterial phytases and uses thereof
 JOURNAL Patent: WO 0071728-A 1 30-NOV-2000;
 Diversa Corporation (US)
 FEATURES
 source 1..1323
 /organism="Escherichia coli"
 /mol_type="unassigned DNA"
 /db_xref="taxon:562"
 1..1323
 /note="unnamed protein product"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC21810.1"
 /db_xref="GI:12226904"
 /translation="MKAILIPFLSLILPLTPQSAFAQSEPELKLESVIVSRHGVRA
 TKATOLMDVTPDAMPVPLGLMUTPRGSLIYVGHYORLVAADGLAKGCPQS
 GOVALIADVDERTKTGGEAFAGAPDAITVHOADPSSDPLFNPDKTGCOLDNA
 NTPYALISRAAGSITADPFGHQTAFLRELRANTPQSNLCLKRKQDESCSLTQALPS
 ELKVSADNVSLTGAIVSLASMLTETFLVDAQAGMPGKRTIDSHQNTLISLNAGP
 VLTORTPEVARAPETPLDLITMAALTPIHPKQAYGVTLPTPSVFLIAGHDNTNLNAG
 ALBLNMTLPGQPDNTPPGSELVFERMRLSDNSQMIQVSLVFQTLQWRDXTPLSNT
 PPGRVKTLIACGEBRNAGMCSLAGFTQIVVEARI PACSLSHHHHHH"

ORIGIN


```

QY 121 GTGCGTCTCCAAACCAAGGCGACGCACTGATCAGATGTCACCCCGACGATGGCCA 180
DB 122 GTGCGTCTCCAAACCAAGGCGACGCACTGATCAGATGTCACCCCGACGATGGCCA 180
QY 181 ACTGGCCGGTAAACCTGGGTTGGCTGACACCGCGAGGTGGTGAAGTAAATCCCTATCTC 240
DB 181 ACTGGCCGGTAAACCTGGGTTGGCTGACACCGCGAGGTGGTGAAGTAAATCCCTATCTC 240
QY 241 GGAACATTACCAACGCGAGGCTCTGATAGCGACGGATTCCTGGCGAAAAAGGCTCCCG 300
DB 241 GGAACATTACCAACGCGAGGCTCTGATAGCGACGGATTCCTGGCGAAAAAGGCTCCCG 300
QY 301 CAGTCTGTCAAGTCCGATTTATTTGCTGATGTCAGACGATCCCGTAAACAGGCGAA 360
DB 301 CAGTCTGTCAAGTCCGATTTATTTGCTGATGTCAGACGATCCCGTAAACAGGCGAA 360
QY 361 GCGTTTCGCGCGCGGCTGGCACTGACTGTGCATTAACCGTACATACCAAGGAGATAG 420
DB 361 GCGTTTCGCGCGCGGCTGGCACTGACTGTGCATTAACCGTACATACCAAGGAGATAG 420
QY 421 TCCAGTCCGATCCGTTATTTAATCTCTAAACCTGGGCTTTGCGAATCGATTAACGG 480
DB 421 TCCAGTCCGATCCGTTATTTAATCTCTAAACCTGGGCTTTGCGAATCGATTAACGG 480
QY 481 AACGTGACTGACGCGATCTCTACAGAGGCGAGAGGCTCAATTGCTGACTTTACCGGCGAT 540
DB 481 AACGTGACTGACGCGATCTCTACAGAGGCGAGAGGCTCAATTGCTGACTTTACCGGCGAT 540
QY 541 CGGCAAAACGCGCTTTTCGCGAATGGAACGCGGTCTTAATTTTCGCGAATCAAACTTGTGC 600
DB 541 CGGCAAAACGCGCTTTTCGCGAATGGAACGCGGTCTTAATTTTCGCGAATCAAACTTGTGC 600
QY 601 CTTAAACGTGAACAAACAGACGAAAGCTTTCAATTAACGAGCATTAACATCGGAATC 660
DB 601 CTTAAACGTGAACAAACAGACGAAAGCTTTCAATTAACGAGCATTAACATCGGAATC 660
QY 661 AAGGTAGCGCGCAATGCTCTAATTAACGCGTGCCTGAAGCTCGCATCAATGCGAGC 720
DB 661 AAGGTAGCGCGCAATGCTCTAATTAACGCGTGCCTGAAGCTCGCATCAATGCGAGC 720
QY 721 GAGATATTTCTCTGCAACAAACAGGGAATGCGGAGCCGCGGTGGGAAAGATCAAC 780
DB 721 GAGATATTTCTCTGCAACAAACAGGGAATGCGGAGCCGCGGTGGGAAAGATCAAC 780
QY 781 GATTCAACCAATGGAACCACTTGTCTAATGTTGCATTAACGCGCAATTTTATTTGCTACA 840
DB 781 GATTCAACCAATGGAACCACTTGTCTAATGTTGCATTAACGCGCAATTTTATTTGCTACA 840
QY 841 CGCAGCGCAAGAGTTGCCGCGACGCGCGCACCCCGTTATTGGAATTTGATCAAGCAGCG 900
DB 841 CGCAGCGCAAGAGTTGCCGCGACGCGCGCACCCCGTTATTGGAATTTGATCAAGCAGCG 900
QY 901 TTGACGCGCCATCAACCGCAAAACAGGCGTATGTTGATCAATTAACCACTTCAGTACTG 960
DB 901 TTGACGCGCCATCAACCGCAAAACAGGCGTATGTTGATCAATTAACCACTTCAGTACTG 960
QY 961 TTTATTTGCGCGACAGATTAATTAATGCGCAATTCGCGCGCGCACTGAGCTCAACTGG 1020
DB 961 TTTATTTGCGCGACAGATTAATTAATGCGCAATTCGCGCGCGCACTGAGCTCAACTGG 1020
QY 1021 ACGCTTCCCGGTCAACCGGATTAACAGCGCGCAGGAGGTGAACCTGATTTGAAGCGTGG 1080
DB 1021 ACGCTTCCCGGTCAACCGGATTAACAGCGCGCAGGAGGTGAACCTGATTTGAAGCGTGG 1080
QY 1081 CGTGGCTTAAGGATTAACAGCGAGTGAATTCAGGTTTCGCTGATCTTCAGACCTTAAAG 1140
DB 1081 CGTGGCTTAAGGATTAACAGCGAGTGAATTCAGGTTTCGCTGATCTTCAGACCTTAAAG 1140
QY 1141 CAGATGCGGATTAACAGCGCGCTGTCTAATTAATGCGCGCGCGGAGAGGTGAATCTGAC 1200
DB 1141 CAGATGCGGATTAACAGCGCGCTGTCTAATTAATGCGCGCGCGGAGAGGTGAATCTGAC 1200
QY 1201 CTGGCAGATGTGAAGAGGAATGCGAGGCGATGTGTGTTGAGCAGGTTTACGCA 1260

```

```

DB 1201 CTGGCAGATGTGAAGAGGAATGCGAGGCGATGTGTGTTGAGCAGGTTTACGCA 1260
QY 1261 ATCTGGAATGAAGACGCACTACCGGCTGACAGTTTGAAGTCTCATCAATCAATCAAC 1320
DB 1261 ATCTGGAATGAAGACGCACTACCGGCTGACAGTTTGAAGTCTCATCAATCAATCAAC 1320
QY 1321 TAA 1323
DB 1321 TAA 1323

RESULT 8
AX356566 1323 bp DNA linear PAT 06-FEB-2002
LOCUS Sequence 1 from Patent WO0190333.
DEFINITION AX356566
ACCESSION AX356566
VERSION AX356566.1 GI:18621053
KEYWORDS
SOURCE
ORGANISM
Escherichia coli
Bacterium; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE
1 Short, J.M., Kretz, K.A., Gray, K.A., Barton, N.R., Garrett, J.B. and
O'Donoghue, B.
Recombinant Bacterial phytases and uses thereof
Patent: WO 0190333-A 1 29-NOV-2001;
JOURNAL DIVERSA CORPORATION (US)
FEATURES
source
1..1323
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"
1..1323
/note="unnamed protein product"
/codon_start=1
/translation="MKALILPFLSLILPLTPQSAFAQSEPELKESVIVISNRHGRAP
TKATQIMODVTPDAMPVPLKGLMTPRGGELIYVIGHYORBLVADGLLKKGROS
GOVAILIDVDRRTKRTGAPAGALPDALITVHQDAPDSDDPLPNPKTVYCOLDMA
NTVDAILSRAGSIALDPFGHQTQFRLERLANTPOSUCLIKREKODESCUTOLNP
ELKVSADNVSLTGAIVSLASMLTEIPLLDLQAGQMPGPGWRTDSDQWNTLSLNAQF
YLQRTPEVARSRAPIPLDLIYMAALTPPPOKQYGVLPVSLVFIAGHDNLTNLIG
ALEINMTLPQDPDTPPGSELVFERMRSLSNSQMIQVSLVFOTLQKRDTPSLNT
PFGVKLTLACBERNAQMSLAFITQIVNEARIPACSLASHHHHH"
ORIGIN
Query Match 100.0%; Score 1323; DB 2; Length 1323;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAGGAGCTTAATCCATTTTATCTCTTCGATTCCTGTTAACCCCGGAATCTCA 60
DB 1 ATGAAGGAGCTTAATCCATTTTATCTCTTCGATTCCTGTTAACCCCGGAATCTCA 60
QY 61 TTCGCTCAGATGAGCCGGAAGCTGAAGTGAAGTGTGATTTGTCAGTCTGATGGT 120
DB 61 TTCGCTCAGATGAGCCGGAAGCTGAAGTGAAGTGTGATTTGTCAGTCTGATGGT 120
QY 121 GTGCGTCTCCAAACCAAGGCGACGCACTGATGAGATGTCAACCCGACGATGGCCA 180
DB 121 GTGCGTCTCCAAACCAAGGCGACGCACTGATGAGATGTCAACCCGACGATGGCCA 180
QY 181 AACTGCGCGGTAACCTGGGTTGGCTGACACCGCGAGGTGGTGAAGTAAATCCCTATCTC 240
DB 181 AACTGCGCGGTAACCTGGGTTGGCTGACACCGCGAGGTGGTGAAGTAAATCCCTATCTC 240
QY 241 GGAACATTACCAACGCGAGGCTCTGATAGCGACGGATTCCTGGCGAAAAAGGCTCCCG 300
DB 241 GGAACATTACCAACGCGAGGCTCTGATAGCGACGGATTCCTGGCGAAAAAGGCTCCCG 300

```

QY 301 CAGTCTGGTCAAGTCCGCGATTATTGCTGATGTCAGACGACGCGTACCCTTTAAACAGGCGAA 360
DB 301 CAGTCTGGTCAAGTCCGCGATTATTGCTGATGTCAGACGACGCGTACCCTTTAAACAGGCGAA 360
QY 361 GCCTTCGCGCGCGGCTTGCGCACTGATCTGTGCATTAACCGTACATACCAGGAGATACG 420
DB 361 GCCTTCGCGCGCGGCTTGCGCACTGATCTGTGCATTAACCGTACATACCAGGAGATACG 420
QY 421 TCCAGTCCCGATCCGCTTATTATCTCTTAAATACTGCGCTTTGCGCACTGATACCGG 480
DB 421 TCCAGTCCCGATCCGCTTATTATCTCTTAAATACTGCGCTTTGCGCACTGATACCGG 480
QY 481 AACGCTGACGACGCGATCTCTGACGAGGAGGAGGCTCAATTGCTGACTTTAACGGGAGAT 540
DB 481 AACGCTGACGACGCGATCTCTGACGAGGAGGAGGCTCAATTGCTGACTTTAACGGGAGAT 540
QY 541 CGGCAAAAGGGGTTTCGCGAACTGGAACGGGTGCTTAAATTTTCGCAATCAAACTTGTGC 600
DB 541 CGGCAAAAGGGGTTTCGCGAACTGGAACGGGTGCTTAAATTTTCGCAATCAAACTTGTGC 600
QY 601 CTTAAACGTGAGAAACAGAGCGAAAGCTGTTCAATTAACGACGAGCATTAACCATCGAATC 660
DB 601 CTTAAACGTGAGAAACAGAGCGAAAGCTGTTCAATTAACGACGAGCATTAACCATCGAATC 660
QY 661 AAGGTGAGCGCGGCAATGTCTCATTTAACCGGTGCGGTAAAGCTGCAATCAATGCTGACG 720
DB 661 AAGGTGAGCGCGGCAATGTCTCATTTAACCGGTGCGGTAAAGCTGCAATCAATGCTGACG 720
QY 721 GAGATATTTCTCTCGCAACAGCAGAGGAATGCGAGCGCGGGTGGGGAAGATCAAC 780
DB 721 GAGATATTTCTCTCGCAACAGCAGAGGAATGCGAGCGCGGGTGGGGAAGATCAAC 780
QY 781 GATTCAACACGATGGAACACCTTGTGAAGTTGCTAAACGGCATTTTATTTGCTAAC 840
DB 781 GATTCAACACGATGGAACACCTTGTGAAGTTGCTAAACGGCATTTTATTTGCTAAC 840
QY 841 CGCAGCGCAGAGGTTGCGCGACGCGCGCACCCGTTATTGATTTGATCACTGACGCG 900
DB 841 CGCAGCGCAGAGGTTGCGCGACGCGCGCACCCGTTATTGATTTGATCACTGACGCG 900
QY 901 TTGACGCGCCCATCAACCGCAAAACAGCGGTATGCTGACATTAACCATTCAGTACTG 960
DB 901 TTGACGCGCCCATCAACCGCAAAACAGCGGTATGCTGACATTAACCATTCAGTACTG 960
QY 961 TTTATTTGCGGACAGATTAATCTGTGGGAATCTCGGCGGCGCACTGAGCTCAACTGG 1020
DB 961 TTTATTTGCGGACAGATTAATCTGTGGGAATCTCGGCGGCGCACTGAGCTCAACTGG 1020
QY 1021 ACGCTTCCCGGTCAAGCGGATTAACAGCGCGCAGGTTGTAATCTGTTGTAACGCTGG 1080
DB 1021 ACGCTTCCCGGTCAAGCGGATTAACAGCGCGCAGGTTGTAATCTGTTGTAACGCTGG 1080
QY 1081 CGTCGCGCTTAAGGATTAACAGCGGATTAAGGTTGCTGCTGCTTTCAGACTTAAACG 1140
DB 1081 CGTCGCGCTTAAGGATTAACAGCGGATTAAGGTTGCTGCTGCTTTCAGACTTAAACG 1140
QY 1141 CAGATGCGGTATTAACAGCGCGCTGTCTATTAATACGCGCGCGAGAGGTGAACCTGAC 1200
DB 1141 CAGATGCGGTATTAACAGCGCGCTGTCTATTAATACGCGCGCGAGAGGTGAACCTGAC 1200
QY 1201 CTGGCAGAGTGAAGAGGAAATGCGAGGCGATGTGTTCTGTTGCGAGGTTTACGCA 1260
DB 1201 CTGGCAGAGTGAAGAGGAAATGCGAGGCGATGTGTTCTGTTGCGAGGTTTACGCA 1260
QY 1261 ATCTGGAATGAAGACGCAATACCGGCTGTCAGTTGAGATCTCATCAATCAATCAATC 1320
DB 1261 ATCTGGAATGAAGACGCAATACCGGCTGTCAGTTGAGATCTCATCAATCAATCAATC 1320
QY 1321 TAA 1323
DB 1321 TAA 1323

RESULT 9
AR636190
LOCUS AR636190 1901 bp DNA linear PAT 20-APR-2005
DEFINITION Sequence 7 from patent US 6855365.
ACCESSION AR636190
VERSION AR636190.1 GI:62768873
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE
AUTHORS Short,J.M., Krezs,K.A., Gray,K.A., Barton,N.R., Garrett,J.B.,
O'Donoghue,B. and Mathur,E.J.
TITLE Recombinant bacterial phycases and uses thereof
JOURNAL Patent: US 6855365-A 7 15-FEB-2005;
Diversa Corporation; San Diego, CA
FEATURES
source
1..1901
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 97.4%; Score 1287.8; DB 2; Length 1901;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1293; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 ATGAAAGGATCTTAATCCATTTTATCTCTGATTCGGTTAACCCGCAATCTGCA 60
DB 188 ATGAAAGGATCTTAATCCATTTTATCTCTGATTCGGTTAACCCGCAATCTGCA 247
QY 61 TTCCTCAGATGAGCCGAGACTGAGAGTGAAGTGTGATGTCAGTCTGATGCT 120
DB 248 TTCCTCAGATGAGCCGAGACTGAGAGTGAAGTGTGATGTCAGTCTGATGCT 307
QY 121 GTGCGTCTCCACCAAGGCCACGCACTGATGCAAGATGTCACCCGACGATGGCCA 180
DB 308 GTGCGTCTCCACCAAGGCCACGCACTGATGCAAGATGTCACCCGACGATGGCCA 367
QY 181 ACGTCCCGGTAAACCTGGGTGCTGACACCGGCGGTGAGCTAATCGCTATCTC 240
DB 368 ACGTCCCGGTAAACCTGGGTGCTGACACCGGCGGTGAGCTAATCGCTATCTC 427
QY 241 GGAATTAACCAACGACGCGTCTGATGCTGATGCTGATGCTGATGCTGATGCT 300
DB 428 GGAATTAACCAACGACGCGTCTGATGCTGATGCTGATGCTGATGCTGATGCT 487
QY 301 CAGTCTGTCAGTCTGATTTGCTGATGCTGATGCTGATGCTGATGCTGATGCT 360
DB 488 CAGTCTGTCAGTCTGATTTGCTGATGCTGATGCTGATGCTGATGCTGATGCT 547
QY 361 GCCTTCGCGCGCGGCTTGCGCACTGATCTGTGCAATTAACCGTACATACCAGGATACG 420
DB 548 GCCTTCGCGCGCGGCTTGCGCACTGATCTGTGCAATTAACCGTACATACCAGGATACG 607
QY 421 TCCAGTCCCGATCCGCTTATTATCTCTTAAATACTGCGCTTTGCGCACTGATACCGG 480
DB 608 TCCAGTCCCGATCCGCTTATTATCTCTTAAATACTGCGCTTTGCGCACTGATACCGG 667
QY 481 AACGCTGACGACGCGATCTCTGACGAGGAGGAGGCTCAATTGCTGACTTTAACGGGAGAT 540
DB 668 AACGCTGACGACGCGATCTCTGACGAGGAGGAGGCTCAATTGCTGACTTTAACGGGAGAT 727
QY 541 CGGCAAAAGGGGTTTCGCGAACTGGAACGGGTGCTTAAATTTTCGCAATCAAACTTGTGC 600
DB 728 CGGCAAAAGGGGTTTCGCGAACTGGAACGGGTGCTTAAATTTTCGCAATCAAACTTGTGC 787
QY 601 CTTAAACGTGAGAAACAGAGCGAAAGCTGTTCAATTAACGACGAGCATTAACCATCGAATC 660
DB 788 CTTAAACGTGAGAAACAGAGCGAAAGCTGTTCAATTAACGACGAGCATTAACCATCGAATC 847
QY 661 AAGGTGAGCGCGGCAATGTCTCATTTAACCGGTGCGGTAAAGCTGCAATCAATGCTGACG 720

Db 848 AAGGTAGCGCCGACATATGTCTGATTAAACCGGTGCGGTAAAGCTTCGATCATATGCTGACG 907
 QY 721 GAGATATTTCTCTGCAACAGGCAATGCGGAAATGCGGAGCGCGGGTGGGAAAGATACAC 780
 Db 908 GAAATATTTCTCTGCAACAGGCAATGCGGAAATGCGGAGCGCGGGTGGGAAAGATACAC 967
 QY 781 GATTCAACCAAGTGAACAACCTTGTAAAGTTGCAATTAACGCAATTTTATTTGCTACAA 840
 Db 968 GATTCAACCAAGTGAACAACCTTGTAAAGTTGCAATTAACGCAATTTTATTTGCTACAA 1027
 QY 841 CGCACCCGCAAGAGTTGCGCGACCGCGGCAACCCCGTTATTTGATTTGATCATGCAACG 900
 Db 1028 CGCACCCGCAAGAGTTGCGCGACCGCGGCAACCCCGTTATTTGATTTGATCATGCAACG 1087
 QY 901 TTGAGCGCCCATTCACCGCAAAAACAGCGGTATGATGATTAATCCCACTTCAATCACTG 960
 Db 1088 TTGAGCGCCCATTCACCGCAAAAACAGCGGTATGATGATTAATCCCACTTCAATCACTG 1147
 QY 961 TTTATTTGCGGACACAGTACTAATCTGCAAAATCTGCGCGGCGCACTGAGCTCAACTG 1020
 Db 1148 TTTATTTGCGGACACAGTACTAATCTGCAAAATCTGCGCGGCGCACTGAGCTCAACTG 1207
 QY 1021 AGCGTTCCCGGTGACCGCGATTAACAGCGCGCAAGTGTGAATCTGTTGTGAACGCTG 1080
 Db 1208 AGCGTTCCCGGTGACCGCGATTAACAGCGCGCAAGTGTGAATCTGTTGTGAACGCTG 1267
 QY 1081 CGTCCGCTTAAGCATTAACAGCGCGATTAATCAAGTTTCTGCTGCTTCCAGACTTTACG 1140
 Db 1268 CGTCCGCTTAAGCATTAACAGCGCGATTAATCAAGTTTCTGCTGCTTCCAGACTTTACG 1327
 QY 1141 CAGATCGTGAATTAACAGCGCGCTGTCTAATTAATACCGCGCGGAGGAGGAACTGAC 1200
 Db 1328 CAGATCGTGAATTAACAGCGCGCTGTCTAATTAATACCGCGCGGAGGAGGAACTGAC 1387
 QY 1201 CTGGCAGATGTGAAGAGGAAATGCGCAGGCGATGTTCTGTTGCAAGTTTACGCA 1260
 Db 1388 CTGGCAGATGTGAAGAGGAAATGCGCAGGCGATGTTCTGTTGCAAGTTTACGCA 1447
 QY 1261 ATCGTAATGAAGCAGCATACCGGCGTGTGCACTTTGAGAT 1300
 Db 1448 ATCGTAATGAAGCAGCATACCGGCGTGTGCACTTTGAGAT 1487
 RESULT 10
 AX356572
 LOCUS AX356572 1901 bp DNA linear PAT 06-FEB-2002
 DEFINITION Sequence 7 from Patent WO0190333.
 ACCESSION AX356572
 VERSION AX356572.1 GI:18621059
 KEYWORDS
 SOURCE Escherichia coli
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.
 REFERENCE 1 Short, J.M., Kretz, K.A., Gray, K.A., Batton, N.R., Garrett, J.B. and O'Donoghue, E. Recombinant bacterial phytases and uses thereof Patent: WO 0190333-A 7 29-NOV-2001;
 JOURNAL DIVERSA CORPORATION (US)
 FEATURES Location/Qualifiers
 source 1..1901
 /organism="Escherichia coli"
 /mol_type="unassigned DNA"
 /db_xref="taxon:562"
 ORIGIN
 Query Match 97.4%; Score 1287.8; DB 2; Length 1901;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1293; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 1 ATGAAGCGATCTTAATCCCATTTTATCTTCTGATTCGTTAACCCGCAATCTGCA 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db

Db 188 ATGAAGCGATCTTAATCCCATTTTATCTTCTGATTCGTTAACCCGCAATCTGCA 247
 QY 61 TTGCTCAAGATGAGCGCGAGTGAAGTGAAGTGTGTGATTTGTCACTGCTCATGCT 120
 Db 248 TTGCTCAAGATGAGCGCGAGTGAAGTGAAGTGTGTGATTTGTCACTGCTCATGCT 307
 QY 121 GTGCGTCTCAACCAAGGCAACGCAATGATGACAGATGTCAACCCGACGATGAGCA 180
 Db 308 GTGCGTCTCAACCAAGGCAACGCAATGATGACAGATGTCAACCCGACGATGAGCA 367
 QY 181 ACCTGCGCGATTAACCTGAGTTGCTGACACCGCGAGTGTGATGATTAATCCCTATCTC 240
 Db 368 ACCTGCGCGATTAACCTGAGTTGCTGACACCGCGAGTGTGATGATTAATCCCTATCTC 427
 QY 241 GGAATTAACCAACGCAAGTGTGTGATGCTGACGATTTGCTGCGAAAAAGGCTCGCG 300
 Db 428 GGAATTAACCAACGCAAGTGTGTGATGCTGACGATTTGCTGCGAAAAAGGCTCGCG 487
 QY 301 CAGTCTGTGACAGTGTGCAATTAATGCTGATGTCGATGAGAGCGTAAACCGCGAAG 360
 Db 488 CAGTCTGTGACAGTGTGCAATTAATGCTGATGTCGATGAGAGCGTAAACCGCGAAG 547
 QY 481 AACGTGACGAGCGCATCTCTGACGAGGAGGAGGATCAATTGCTGACCTTACCGGCGCAT 540
 Db 608 TCACGTCGCGATCCGTTATTTAATCCTTAATAACGCGGTTTGCACATGAGTAAACGCG 667
 QY 421 TCCAGTCCCGATCCGTTATTTAATCCTTAATAACGCGGTTTGCACATGAGTAAACGCG 480
 Db 608 TCACGTCGCGATCCGTTATTTAATCCTTAATAACGCGGTTTGCACATGAGTAAACGCG 667
 QY 481 AACGTGACGAGCGCATCTCTGACGAGGAGGAGGATCAATTGCTGACCTTACCGGCGCAT 540
 Db 668 AACGTGACGAGCGCATCTCTGACGAGGAGGAGGATCAATTGCTGACCTTACCGGCGCAT 727
 QY 541 CGGCAAAACGCGCTTTGCGCAATGGAACGCGGTGCTTAATTTTCCGCAATCAACTGTGCG 600
 Db 728 CGGCAAAACGCGCTTTGCGCAATGGAACGCGGTGCTTAATTTTCCGCAATCAACTGTGCG 787
 QY 601 CTTTAAACGTGAAGAACAGGACGAAGCTGTTCAATTACGACGATTAACATCGAGAACTC 660
 Db 788 CTTTAAACGTGAAGAACAGGACGAAGCTGTTCAATTACGACGATTAACATCGAGAACTC 847
 QY 661 AAGGTAGCGCGGACAAATGCTCATTAACCGGTGCGGTAAAGCTGCAATCAATGCTGAG 720
 Db 848 AAGGTAGCGCGGACAAATGCTCATTAACCGGTGCGGTAAAGCTGCAATCAATGCTGAG 907
 QY 721 GAGATATTTCTCTGCAACAGGCAATGCGGAAATGCGGAGCGGGTGGGAAAGATCAC 780
 Db 908 GAGATATTTCTCTGCAACAGGCAATGCGGAAATGCGGAGCGGGTGGGAAAGATCAC 967
 QY 781 GATTCAACCAAGTGAACAACCTTGTAAAGTTGCAATTAACGCAATTTTATTTGCTACAA 840
 Db 968 GATTCAACCAAGTGAACAACCTTGTAAAGTTGCAATTAACGCAATTTTATTTGCTACAA 1027
 QY 841 CGCACCCGCAAGAGTTGCGCGACCGCGGCAACCCCGTTATTTGATTTGATCATGCAACG 900
 Db 1028 CGCACCCGCAAGAGTTGCGCGACCGCGGCAACCCCGTTATTTGATTTGATCATGCAACG 1087
 QY 901 TTGAGCGCCCATTCACCGCAAAAACAGCGGTATGATGATTAATCCCACTTCAATCACTG 960
 Db 1088 TTGAGCGCCCATTCACCGCAAAAACAGCGGTATGATGATTAATCCCACTTCAATCACTG 1147
 QY 961 TTTATTTGCGGACACAGTACTAATCTGCAAAATCTGCGCGGCGCACTGAGCTCAACTG 1020
 Db 1148 TTTATTTGCGGACACAGTACTAATCTGCAAAATCTGCGCGGCGCACTGAGCTCAACTG 1207
 QY 1021 AGCGTTCCCGGTGACCGCGATTAACAGCGCGCAAGTGTGAATCTGTTGTGAACGCTG 1080
 Db 1208 AGCGTTCCCGGTGACCGCGATTAACAGCGCGCAAGTGTGAATCTGTTGTGAACGCTG 1267
 QY 1081 CGTCCGCTTAAGCATTAACAGCGCGATTAATCAAGTTTCTGCTGCTTCCAGACTTTACG 1140
 Db 1268 CGTCCGCTTAAGCATTAACAGCGCGATTAATCAAGTTTCTGCTGCTTCCAGACTTTACG 1327

[illegible]

		/gene="appa"	/standard_name="acid phosphatase"
		/codon_start=1	/transl_table=11
		/product="periplasmic phosphoanhydride phosphohydrolase"	/protein_id="AA72086.1"
		/db_xref="GI:145285"	/translation="MKAILLPFSLPLPTPQSAFAQSPEPKLESVVIYSRHGVARR TKAIQLMDVTDPDMFTPWPMVPLGMLTRGGELIAYLDHYQRRLVAAGLLAKKCPOS GOVAIIADVDERTKTEBAFAMGLAPCALIVHQADTISPDLEFNPLKTGCQLDANA NVTDAISRAGGSIADEPTGHROZAFRELERIANPOSNLCIKRKOPESCSITOLAES EIKASANSVSLTGAVLSAASMTETIELQAQGWEPGMGRITDSHONWTLISLANOGE YLLORTBEVASRSATRPILLDLIKTALTPHPKXKAVGYLIPSVLYFIAGHDINLIANG ALEHNMTLPQGPNDTPPGGELVFERRRSLSDNSGMIVSYLVFOTLQQMRDKTPIASLTNT PPGEVKTLTAGCEERNAQMCISLAGTFQIVNEARI PACSL"
	sig_peptide	188..253	
	mat_peptide	254..1483	
		/gene="appa"	/standard_name="R92A mutant"
		/product="periplasmic phosphoanhydride phosphohydrolase"	/citation=[2]
	variation	299..301	
		/gene="appa"	/standard_name="R16A mutant"
		/note="created by site directed mutagenesis"	/citation=[2]
	variation	299..301	
		/gene="appa"	/standard_name="R20A mutant"
		/note="created by site directed mutagenesis"	/citation=[2]
	variation	299..301	
		/gene="appa"	/standard_name="H17N mutant"
		/note="Created by site directed mutagenesis"	/citation=[2]
	variation	302..304	
		/gene="appa"	/standard_name="D304A mutant"
		/note="Created by site directed mutagenesis"	/citation=[3]
	variation	1163..1165	
		/gene="appa"	/standard_name="D304F mutant"
		/note="Created by site directed mutagenesis"	/citation=[3]
	variation	1163..1165	
		/phenotype="Asp 304 changed to Ala"	/replace="gct"
	variation	1163..1165	
		/gene="appa"	/standard_name="D304E mutant"
		/note="Created by site directed mutagenesis"	/citation=[3]
	terminator	1497..1522	
		/gene="appa"	/note="rho-independent terminator"
ORIGIN			
Query Match	97.4%; Score 1287.8; DB 15; Length 1901;		
Best Local Similarity	99.5%; Pred. No. 0;		
Matches 1293; Conservative	0; Mismatches 7; Indels 0; Gaps 0;		
61 ATGAAGGAGCTTAATCCCATTTTATCTCTTGATTCGGTAAACCGCAGATCTGCA	60		
Db	188 ATGAAGGAGCTTAATCCCATTTTATCTCTTGATTCGGTAAACCGCAGATCTGCA	247	
61 TTTCGTCAAGTGAACCGCAGCTGAAGCTGAAAATGTGTGATTGTCAATTCATCATGT	120		

```

Db      248 TTGCGTCAGAGTGAAGCCGAGCTGAAGCTGGAAGAGTGGTGAATTCAGTGGTCAATGGT 307
Qy      121 GGGCGGCTCCCAACCAAGCCCAAGCACTGATGCAAGAGATGTCACCCCAAGCAATGGCCA 180
Db      308 GTGCGGTCTCAACCAAGGCCAAGCACTGATGCAAGAGATGTCACCCCAAGCAATGGCCA 367
Qy      181 ACCTGGCCGGTAAACTGGGTTGGCTGACACCGGAGGAGTGAAGCTAAATCCGCTATCTC 240
Db      368 ACTGGCCGGTAAACTGGGTTGGCTGACACCGGAGGAGTGAAGCTAAATCCGCTATCTC 427
Qy      241 GGAATTATCAACAGCCAGCGTCTGTAGCCGACGAAATGCTGGCAAAAAAGGCTGCCG 300
Db      428 GGAATTATCAACAGCCAGCGTCTGTAGCCGACGAAATGCTGGCAAAAAAGGCTGCCG 487
Qy      301 CAGTCTGGTCAAGTCCGCAATTAATGCTGATGTCAGAGCGTACCCGTTAAACAGCGCA 360
Db      488 CAGTCTGGTCAAGTCCGCAATTAATGCTGATGTCAGAGCGTACCCGTTAAACAGCGCA 547
Qy      361 GCCTTGGCCGCGGGCTGGCACTGACTGCAATTAACGTAATCCAGAGCAATAG 420
Db      548 GCCTTGGCCGCGGGCTGGCACTGACTGCAATTAACGTAATCCAGAGCAATAG 607
Qy      421 TCCAGTCCCGATCCGTTAATTCCTTAAACCTGGCTTGGCACTGATAACGG 480
Db      608 TCCAGTCCCGATCCGTTAATTCCTTAAACCTGGCTTGGCACTGATAACGG 667
Qy      481 AACGTACTGAAGCGATCTCAGACAGGCAAGAGGTCATATGCTGATTAACGGGCAT 540
Db      668 AACGTACTGAAGCGATCTCAGACAGGCAAGAGGTCATATGCTGATTAACGGGCAT 727
Qy      541 CGGCAAAAGCGGTTCCGCAACTGGAACGGGCTTAATTTCCGCAATCAAACTTGTGC 600
Db      728 CGGCAAAAGCGGTTCCGCAACTGGAACGGGCTTAATTTCCGCAATCAAACTTGTGC 787
Qy      601 CTTAAACGTGAGAAACAGGACGAAAGCTTTCATTAACGAGCATTACATCGAACTC 660
Db      788 CTTAAACGTGAGAAACAGGACGAAAGCTTTCATTAACGAGCATTACATCGAACTC 847
Qy      661 AAGGTAGCGCCGACATATCTCTAATTAACCGGTGCGTAAGCTCGCATCAATGCTGAG 720
Db      848 AAGGTAGCGCCGACATATCTCTAATTAACCGGTGCGTAAGCTCGCATCAATGCTGAG 907
Qy      721 GAGATATTTCTCTCGAACAAGCAAGGAAATGCGGAGCCGGGTGGGAAAGATCAAC 780
Db      908 GAGATATTTCTCTCGAACAAGCAAGGAAATGCGGAGCCGGGTGGGAAAGATCAAC 967
Qy      781 GATTCAACAGTGAACACCTTGCTAATGTTGATTAACGCGCAATTTAATTTGCTACA 840
Db      968 GATTCAACAGTGAACACCTTGCTAATGTTGATTAACGCGCAATTTAATTTGCTACA 1027
Qy      841 CGCAGCGCAGAGGTTGCGCGAGCGCGCACCCCGTTATTGGAATTTGATGAGCAGC 900
Db      1028 CGCAGCGCAGAGGTTGCGCGAGCGCGCACCCCGTTATTGGAATTTGATGAGCAGC 1087
Qy      901 TTGAGCGCCCATCCACCGCAAAAACAGGCGTATGCTGATCAATTAACCTTCAAGTCTG 960
Db      1088 TTGAGCGCCCATCCACCGCAAAAACAGGCGTATGCTGATCAATTAACCTTCAAGTCTG 1147
Qy      961 TTTATTTGCGGACAGATCTAATCTGGAATTTCTGCGGGGCACTGAGCTCAACTGG 1020
Db      1148 TTTATTTGCGGACAGATCTAATCTGGAATTTCTGCGGGGCACTGAGCTCAACTGG 1207
Qy      1021 ACGCTTCCCGATCAAGCCGAGTAAACGCGCCAGGTGTGAATCTGTGTTTGAACGCTGG 1080
Db      1208 ACGCTTCCCGATCAAGCCGAGTAAACGCGCCAGGTGTGAATCTGTGTTTGAACGCTGG 1267
Qy      1081 CTTGGGCTTAAGGATTAACGCGAGTGAATTCAGGTTTCCCTGTCTTCAAGCTTTACG 1140
Db      1268 CTTGGGCTTAAGGATTAACGCGAGTGAATTCAGGTTTCCCTGTCTTCAAGCTTTACG 1327
Qy      1141 CAGATGCGTATTAACGCGCGTGTATTAATTAACGCGCGCGGAGAGGTGAACCTGACC 1200

```

```

Db      1328 CAGATGCGTATTAACGCGCGTGTCTATTAATTAACGCGCGGAGAGGTGAACCTGACC 1387
Qy      1201 CTGGCAGATTTGGAAGAGGAATGCGGAGCATGTGTTCTGTGGCAGTTTACGAA 1260
Db      1388 CTGGCAGATTTGGAAGAGGAATGCGGAGCATGTGTTCTGTGGCAGTTTACGAA 1447
Qy      1261 ATCGTAATGAAGCAGCATACCGGCGTCAAGTTGAGAT 1300
Db      1448 ATCGTAATGAAGCAGCATACCGGCGTCAAGTTGAT 1487

RESULT 12
U00096_10
WPCOMMENT
Sequence split into 47 fragments
Fragment Name      Begin      End
U00096_00          1      110000
U00096_01          10001      21000
U00096_02          20001      31000
U00096_03          30001      41000
U00096_04          40001      51000
U00096_05          50001      61000
U00096_06          60001      71000
U00096_07          70001      81000
U00096_08          80001      91000
U00096_09          90001     101000
U00096_10         100001     111000
U00096_11         110001     121000
U00096_12         120001     131000
U00096_13         130001     141000
U00096_14         140001     151000
U00096_15         150001     161000
U00096_16         160001     171000
U00096_17         170001     181000
U00096_18         180001     191000
U00096_19         190001     201000
U00096_20         200001     211000
U00096_21         210001     221000
U00096_22         220001     231000
U00096_23         230001     241000
U00096_24         240001     251000
U00096_25         250001     261000
U00096_26         260001     271000
U00096_27         270001     281000
U00096_28         280001     291000
U00096_29         290001     301000
U00096_30         300001     311000
U00096_31         310001     321000
U00096_32         320001     331000
U00096_33         330001     341000
U00096_34         340001     351000
U00096_35         350001     361000
U00096_36         360001     371000
U00096_37         370001     381000
U00096_38         380001     391000
U00096_39         390001     401000
U00096_40         400001     411000
U00096_41         410001     421000
U00096_42         420001     431000
U00096_43         430001     441000
U00096_44         440001     451000
U00096_45         450001     461000
U00096_46         460001     4639675
Continuation (11 of 47) of U00096 from base 1000001 (U00096 Escherichia coli K-12 MG1655

Query Match      97.4%; Score 1287.8; DB 15; Length 110000;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1292; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy      1 ATGAAGAGATCTTAATCCATTTTATCTCTGATTCGTTAACCCCGCAATCTCA 60
Db      39840 ATGAAGAGATCTTAATCCATTTTATCTCTGATTCGTTAACCCCGCAATCTCA 39899

```

Qy 61 TTCCGTCAAGTGTAGCCGAGCTGAGAACTGTGATGTTGATGATGTCATGAT 120
Db 3990 TTTCCGTCAAGTGTAGCCGAGCTGAGAACTGTGATGTTGATGATGTCATGATG 39959
Qy 121 GTGGTGTCTCCAACTGAAGCCGCACTGATGCAAGATGTCACTCCGAGCGCATG 180
Db 39960 GTGGTGTCTCCAACTGAAGCCGCACTGATGCAAGATGTCACTCCGAGCGCATG 40019
Qy 181 ACTGGCCGGTAAACTGGGTGGCTGACCGCGGNGGTGGTGGTGGTGGTGGTGGT 240
Db 40020 ACTGGCCGGTAAACTGGGTGGCTGACCGCGGNGGTGGTGGTGGTGGTGGTGGT 40079
Qy 241 GACATTAACAACGCGCAGCGTCTGTAGCCGACGATTTGCTGGCCGAAAAAGGCTGCCG 300
Db 40080 GACATTAACAACGCGCAGCGTCTGTAGCCGACGATTTGCTGGCCGAAAAAGGCTGCCG 40139
Qy 301 CAGTCTGTGTCAGGTGCGATTTATTTGCTGTGATGTCAGACGCTTACCCGTTAAACAGGCGAA 360
Db 40140 CAGTCTGTGTCAGGTGCGATTTATTTGCTGTGATGTCAGACGCTTACCCGTTAAACAGGCGAA 40199
Qy 361 GCTTCCGCGCGCGGCTGGCACTGTGCAATTAACGTAATTAACGAGGCAATTAACG 420
Db 40200 GCTTCCGCGCGCGGCTGGCACTGTGCAATTAACGTAATTAACGAGGCAATTAACG 40259
Qy 421 TCCAGTCCCGATCCGTTATTTATCTCTTAAACTGGCGTTGGCACTGGATTAACGCG 480
Db 40260 TCCAGTCCCGATCCGTTATTTATCTCTTAAACTGGCGTTGGCACTGGATTAACGCG 40319
Qy 481 AACGTACTGACGCGATCTCTGACGAGGCGAGGAGGTCAATTTGCTGACTTTACCGGGCAT 540
Db 40320 AACGTACTGACGCGATCTCTGACGAGGCGAGGAGGTCAATTTGCTGACTTTACCGGGCAT 40379
Qy 541 CGGCAAAAGGGGTTTCGGCACTGGAACGGGNGCTTAATTTTCGCGCATCAAACTTGTGC 600
Db 40380 CGGCAAAAGGGGTTTCGGCACTGGAACGGGNGCTTAATTTTCGCGCATCAAACTTGTGC 40439
Qy 601 CTTAAACGTGAGAAACAGACCAAGCTGTTCAATTAACGAGGCAATTAACGAGACTC 660
Db 40440 CTTAAACGTGAGAAACAGACCAAGCTGTTCAATTAACGAGGCAATTAACGAGACTC 40499
Qy 661 AAGGTAGGCGCCGACATGTCATTAACCGGTGGGTGAAGCTTGGCATGCTGACG 720
Db 40500 AAGGTAGGCGCCGACATGTCATTAACCGGTGGGTGAAGCTTGGCATGCTGACG 40559
Qy 721 GAGATATTTCTCTGCAACAAGCAGAGGAAATGCGAGCGGGTGGGAAAGATCAC 780
Db 40560 GAGATATTTCTCTGCAACAAGCAGAGGAAATGCGAGCGGGTGGGAAAGATCAC 40619
Qy 781 GATTCAACAACAGTGAACAACCTTGTGAATTTGATTAACGCGCAATTTATTTGCTACA 840
Db 40620 GATTCAACAACAGTGAACAACCTTGTGAATTTGATTAACGCGCAATTTATTTGCTACA 40679
Qy 841 CGCAGCGCAAGAGTTGCCCGCAGCGCGCACCCCGTTATTTGATTTGATTCATGCAAGC 900
Db 40680 CGCAGCGCAAGAGTTGCCCGCAGCGCGCACCCCGTTATTTGATTTGATTCATGCAAGC 40739
Qy 901 TTGAAGCCCATCAACCCGCAAAACAGAGGTATGATGATTAACCTTCACTTCACTAC 960
Db 40740 TTGAAGCCCATCAACCCGCAAAACAGAGGTATGATGATTAACCTTCACTTCACTAC 40799
Qy 961 TTTATTTGCGGACAGATTAATCTAATCTGCAAACTTCGCGCGGCACTGAGGCTCAACTG 1020
Db 40800 TTTATTTGCGGACAGATTAATCTAATCTGCAAACTTCGCGCGGCACTGAGGCTCAACTG 40859
Qy 1021 ACGCTTCCCGGTACCGCGGATTAACGCGCGCAGGTGTGAATCTGAGTTTGAACGCTG 1080
Db 40860 ACGCTTCCCGGTACCGCGGATTAACGCGCGCAGGTGTGAATCTGAGTTTGAACGCTG 40919
Qy 1081 CGTGGCTTAAGGATTAACGCGAGGTGATTAAGTTTCTGCTGCTTCCAGACTTACG 1140
Db 40920 CGTGGCTTAAGGATTAACGCGAGGTGATTAAGTTTCTGCTGCTTCCAGACTTACG 40979
Qy 1141 CAGATGCGTATTAACGCGCGCTGTCTATTAAATACGCGCGCGGAGAGGTGAACCTGAC 1200

Db 40980 CAGATGCGTATTAACGCGCGCTGTCTATTAAATACGCGCGCGGAGGTGAACCTGACC 41039
Qy 1201 CTGGCAGGATGTGAAGCGGAAATGCGAGGCGCATGTGTTGTTGGCAGGTTTACGCA 1260
Db 41040 CTGGCAGGATGTGAAGCGGAAATGCGAGGCGCATGTGTTGTTGGCAGGTTTACGCA 41099
Qy 1261 ATCGTAATGAAGCAGCATACCGCGGTGACGTTTGAGAT 1300
Db 41100 ATCGTAATGAAGCAGCATACCGCGGTGACGTTTGAT 41139

RESULT 13
AP009048_10
WPCOMMENT
Sequence split into 47 fragments. LOCUS AP009048 Accession AP009048

Fragment Name	Begin	End
AP009048_00	1	110000
AP009048_01	100001	210000
AP009048_02	200001	310000
AP009048_03	300001	410000
AP009048_04	400001	510000
AP009048_05	500001	610000
AP009048_06	600001	710000
AP009048_07	700001	810000
AP009048_08	800001	910000
AP009048_09	900001	1010000
AP009048_10	1000001	1110000
AP009048_11	1100001	1210000
AP009048_12	1200001	1310000
AP009048_13	1300001	1410000
AP009048_14	1400001	1510000
AP009048_15	1500001	1610000
AP009048_16	1600001	1710000
AP009048_17	1700001	1810000
AP009048_18	1800001	1910000
AP009048_19	1900001	2010000
AP009048_20	2000001	2110000
AP009048_21	2100001	2210000
AP009048_22	2200001	2310000
AP009048_23	2300001	2410000
AP009048_24	2400001	2510000
AP009048_25	2500001	2610000
AP009048_26	2600001	2710000
AP009048_27	2700001	2810000
AP009048_28	2800001	2910000
AP009048_29	2900001	3010000
AP009048_30	3000001	3110000
AP009048_31	3100001	3210000
AP009048_32	3200001	3310000
AP009048_33	3300001	3410000
AP009048_34	3400001	3510000
AP009048_35	3500001	3610000
AP009048_36	3600001	3710000
AP009048_37	3700001	3810000
AP009048_38	3800001	3910000
AP009048_39	3900001	4010000
AP009048_40	4000001	4110000
AP009048_41	4100001	4210000
AP009048_42	4200001	4310000
AP009048_43	4300001	4410000
AP009048_44	4400001	4510000
AP009048_45	4500001	4610000
AP009048_46	4600001	466532

Continuation (11 of 47) of AP009048 from base 1000001 (AP009048 Escherichia coli W3110 D

Query Match 97.4%; Score 1287.8; DB 15; Length 110000;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1292; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 ATGAAGGAGATTTATATCCATTTTATCTTGTGATTCGGTTAAACCCGCAATCTGCA 60
Db 41039 ATGAAGGAGATTTATATCCATTTTATCTTGTGATTCGGTTAAACCCGCAATCTGCA 41098

```

QY 61 TTGCGCTCAGATGAGAGCGGAGCTGGAAGCTGGAAGAGTGTGATTTGTCACTGCTCATGGT 120
DB TTGCGCTCAGATGAGAGCGGAGCTGGAAGAGTGTGATTTGTCACTGCTCATGGT 41158
QY 121 GTGCGGTCTCAACCAAGGCGCAAGCAATGATGACAGATGTCAAGGCGAGCGATGGCCA 180
DB GTGCGGTCTCAACCAAGGCGCAAGCAATGATGACAGATGTCAAGGCGAGCGATGGCCA 41159
QY 181 ACCTGGCCGGTAAACTGAGGTGTGCTGACACCGCGAGGTGTGAGCTAAATCGCTATCTC 240
DB ACCTGGCCGGTAAACTGAGGTGTGCTGACACCGCGAGGTGTGAGCTAAATCGCTATCTC 41219
QY 241 GGAACATTACCAAGCGAGGTCTGTGATGCGGAGATTTGTGCGCAAAAAAGGCTGCCG 300
DB GGAACATTACCAAGCGAGGTCTGTGATGCGGAGATTTGTGCGCAAAAAAGGCTGCCG 41278
QY 301 CAGCTGTGTCAGGTGCGGATTTATGCTGATGTCGACGAGCGTACCCGTAACAGGCGAA 360
DB CAGCTGTGTCAGGTGCGGATTTATGCTGATGTCGACGAGCGTACCCGTAACAGGCGAA 41339
QY 361 GCGTTCGCGCGGCGGCTGACCTGACCTGACATTAACCGTACATTAACCGAGGAGATAG 420
DB GCGTTCGCGCGGCGGCTGACCTGACCTGACATTAACCGTACATTAACCGAGGAGATAG 41399
QY 421 TCCAGTCCCGGATCCGTTATTTATCTCTTAAAACTGGGCTTGGCACTGATTAACGCG 480
DB TCCAGTCCCGGATCCGTTATTTATCTCTTAAAACTGGGCTTGGCACTGATTAACGCG 41459
QY 481 AACGTGATGACGCGATCTCTACGAGGAGGAGGAGGATGCAATTGCTGACTTACCGGGCAT 540
DB AACGTGATGACGCGATCTCTACGAGGAGGAGGAGGATGCAATTGCTGACTTACCGGGCAT 41519
QY 541 CGGCAAAACGGCTTTGCGCAACTGGAACGGGTGCTTAAATTTTCCGCAATCAAACTTGTC 600
DB CGGCAAAACGGCTTTGCGCAACTGGAACGGGTGCTTAAATTTTCCGCAATCAAACTTGTC 41579
QY 601 CTTAAACGTGAAACCAAGACGAAAGCTGTTCAATTAACGAGCATTAACCATGCGAATC 660
DB CTTAAACGTGAAACCAAGACGAAAGCTGTTCAATTAACGAGCATTAACCATGCGAATC 41639
QY 661 AAGGTAGGCGCGCAATGCTCATTAACGGGTGCTGAAAGCTGSCATCAATGCTGAGC 720
DB AAGGTAGGCGCGCAATGCTCATTAACGGGTGCTGAAAGCTGSCATCAATGCTGAGC 41699
QY 721 GAGATATTTCTCTGCAACCAAGCAACGAGATGCGGAGCCGAGGTGCGAAAGATCAAC 780
DB GAGATATTTCTCTGCAACCAAGCAACGAGATGCGGAGCCGAGGTGCGAAAGATCAAC 41759
QY 781 GATTCAACCAAGTGAACACTCTTGCTAGTTGCTAATACGCGCAATTTATTTGCTACA 840
DB GATTCAACCAAGTGAACACTCTTGCTAGTTGCTAATACGCGCAATTTATTTGCTACA 41819
QY 841 CGCAGCGCAGAGGTGCGCGCAGCGCGCAACCGGTATTTGATTTGATCAAGACAGCG 900
DB CGCAGCGCAGAGGTGCGCGCAGCGCGCAACCGGTATTTGATTTGATCAAGACAGCG 41879
QY 901 TTGACGCCCATCAACCGCAAAACAGGCGTATGCTGACATTAACCACTTCAGTACTG 960
DB TTGACGCCCATCAACCGCAAAACAGGCGTATGCTGACATTAACCACTTCAGTACTG 41939
QY 961 TTTATTTGCGGACAGATTAATTAATGCGCAATCTCGCGGCGCACTGAGCTCAACTG 1020
DB TTTATTTGCGGACAGATTAATTAATGCGCAATCTCGCGGCGCACTGAGCTCAACTG 41999
QY 1021 ACGCTTCCCGGTCAACCGGATTAACGCGCGCAGGTGCTGAACTGATTTGAACGCTG 1080
DB ACGCTTCCCGGTCAACCGGATTAACGCGCGCAGGTGCTGAACTGATTTGAACGCTG 42059
QY 1081 CGTGGGCTTAAGGATTAACGCGAGTGAATTCAGGTTTGGCTGCTTCAGACTTTACG 1140
DB CGTGGGCTTAAGGATTAACGCGAGTGAATTCAGGTTTGGCTGCTTCAGACTTTACG 42119
QY 1141 CAGATCCGTGATTAACCGCGCTGCTCATTAATAGCCCGCGGAGAGGTGAACCTGACC 1200
DB CAGATCCGTGATTAACCGCGCTGCTCATTAATAGCCCGCGGAGAGGTGAACCTGACC 42179
QY 1201 CTGGCAGATGTAAGAGGAAATGCGAGGCGATGTGTCGTTGGCAGATTTTACGAA 1260
DB CTGGCAGATGTAAGAGGAAATGCGAGGCGATGTGTCGTTGGCAGATTTTACGAA 42239
QY 1261 ATCGTAATGAAGACGATACCGGCGTGCAGTTGAGAT 1300
DB ATCGTAATGAAGACGATACCGGCGTGCAGTTGAT 42299

```

```

QY 1141 CAGATCCGTGATTAACCGCGCTGCTCATTAATAGCCCGCGGAGAGGTGAACCTGACC 1200
DB 42179 CAGATCCGTGATTAACCGCGCTGCTCATTAATAGCCCGCGGAGAGGTGAACCTGACC 42238
QY 1201 CTGGCAGATGTAAGAGGAAATGCGAGGCGATGTGTCGTTGGCAGATTTTACGAA 1260
DB 42239 CTGGCAGATGTAAGAGGAAATGCGAGGCGATGTGTCGTTGGCAGATTTTACGAA 42298
QY 1261 ATCGTAATGAAGACGATACCGGCGTGCAGTTGAGAT 1300
DB 42299 ATCGTAATGAAGACGATACCGGCGTGCAGTTGAT 42338

RESULT 14
AR636188 1901 bp DNA linear PART 20-APR-2005
LOCUS AR636188
DEFINITION Sequence 5 from patent US 6855365.
ACCESSION AR636188
VERSION AR636188.1 GI:62768871
KEYWORDS
SOURCE
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 1901)
AUTHORS Short, J.M., Kretz, K.A., Gray, K.A., Barton, N.R., Garrett, J.B.,
O'Donoghue, E., and Maltbur, E.J.
TITLE Recombinant Bacterial phylases and uses thereof
JOURNAL Patent: US 6855365-A 5 15-FEB-2005;
Diversa Corporation; San Diego, CA
FEATURES
source
1..1901
location/Qualifiers
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 97.3%; Score 1286.2; DB 2; Length 1901;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1992; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

1 ATGAAGGATCTTATCCATTTTATCTCTTGATTCGTTAACCCCGCAATCTGCA 60
188 ATGAAGGATCTTATCCATTTTATCTCTTGATTCGTTAACCCCGCAATCTGCA 247
QY 61 TTGCTCAGAGTACCGGAGCTGGAAGCTGAAAGTGTGATTTGTGATGTCATGGT 120
DB TTGCTCAGAGTACCGGAGCTGGAAGCTGAAAGTGTGATTTGTGATGTCATGGT 248
QY 121 GTGCGTGTCCAAACCAAGCCCAAGCACTGATGACGATGTCAACCCAGAGCATGGCA 180
DB GTGCGTGTCCAAACCAAGCCCAAGCACTGATGACGATGTCAACCCAGAGCATGGCA 308
QY 308 GTGCGTGTCCAAACCAAGCCCAAGCACTGATGACGATGTCAACCCAGAGCATGGCA 367
DB GTGCGTGTCCAAACCAAGCCCAAGCACTGATGACGATGTCAACCCAGAGCATGGCA 481
QY 481 ACCTGGCGGTAACACTGGTTGGCTGACACCGCGAGGTGTAAGCTATGCTATCTC 240
DB ACCTGGCGGTAACACTGGTTGGCTGACACCGCGAGGTGTAAGCTATGCTATCTC 368
QY 368 ACCTGGCGGTAACACTGGTTGGCTGACACCGCGAGGTGTAAGCTATGCTATCTC 427
DB ACCTGGCGGTAACACTGGTTGGCTGACACCGCGAGGTGTAAGCTATGCTATCTC 421
QY 421 TCCAGTCCCGATCCGTTATTTATTCCTTAAAACTGGCGTTTGCACATGATTAACGCG 480
DB TCCAGTCCCGATCCGTTATTTATTCCTTAAAACTGGCGTTTGCACATGATTAACGCG 608
QY 608 TCCAGTCCCGATCCGTTATTTATTCCTTAAAACTGGCGTTTGCACATGATTAACGCG 667
DB TCCAGTCCCGATCCGTTATTTATTCCTTAAAACTGGCGTTTGCACATGATTAACGCG 481
QY 481 AACGTGATGACGCGATCTCTACGAGGAGGAGGATGCAATTGCTGACTTTACGGGAT 540

```

```
Db      668 AACGGAAGCTGACGGAGTCTGACGACGAGGAGGAGTCAATTGCTGACTTTACGGGGCAT 727
Qy      541 CGGCAACGGCGTTTCGGCAACTGGAAACGGGTGCTTAATTTTCGGCAATCAACTTGTGC 600
Db      728 CGGCAACGGCGTTTCGGCAACTGGAAACGGGTGCTTAATTTTCGGCAATCAACTTGTGC 787
Qy      601 CTTAAACGTGAGAAACAGACGAAAGCTTTCATTAAACGAGGCACTTTCATTCGGAATC 660
Db      788 CTTAAACGTGAGAAACAGACGAAAGCTTTCATTAAACGAGGCACTTTCATTCGGAATC 847
Qy      661 AAGTGAAGCGCCGCAATGTCTCATTTAACGGTGGGTGAAGCTTCATCAATGTCTGACG 720
Db      848 AAGTGAAGCGCCGCAATGTCTCATTTAACGGTGGGTGAAGCTTCATCAATGTCTGACG 907
Qy      721 GAGATATTTCTCTCTGCAACAGCAAGGGAATTCGGAGCCGGGGTGGGGAAGATCAAC 780
Db      908 GAGATATTTCTCTCTGCAACAGCAAGGGAATTCGGAGCCGGGGTGGGGAAGATCAAC 967
Qy      781 GATTCAACACAGTGAACACCTTGCTAGTTTGCAATACGGCAATTTTATTTTGCTACAA 840
Db      968 GATTCAACACAGTGAACACCTTGCTAGTTTGCAATACGGCAATTTTATTTTGCTACAA 1027
Qy      841 CGCAGCGCAGAGTTGCCCGCAGCCGCGCAACCCCGTTATTGATTTGATCAATGCGACG 900
Db      1028 CGCAGCGCAGAGTTGCCCGCAGCCGCGCAACCCCGTTATTGATTTGATCAATGCGACG 1087
Qy      901 TTGAGCGCCCATCCACCGGCAAAAACAGCGCTATGCTGATGACATTAACCACTTCAGTACTG 960
Db      1088 TTGAGCGCCCATCCACCGGCAAAAACAGCGCTATGCTGATGACATTAACCACTTCAGTACTG 1147
Qy      961 TTTATTGCGCGACACAGTACTAATCTGGCAATCTCGCGGCGCACTGAGCTCAACTG 1020
Db      1148 TTTATTGCGCGACACAGTACTAATCTGGCAATCTCGCGGCGCACTGAGCTCAACTG 1207
Qy      1021 ACGCTTCCCGGTCAAGCGGATTAACAGCGCCGCGCAAGTGTGAACCTGTGTGTAACGCTGG 1080
Db      1208 ACGCTTCCCGGTCAAGCGGATTAACAGCGCCGCGCAAGTGTGAACCTGTGTGTAACGCTGG 1267
Qy      1081 CGTGGGCTTAAGCGATTAACAGCGCAATGATTCAGGTTTGGCTGGTCTTTCACACTTAACG 1140
Db      1268 CGTGGGCTTAAGCGATTAACAGCGCAATGATTCAGGTTTGGCTGGTCTTTCACACTTAACG 1327
Qy      1141 CAGATCGCTGATTAACAGCGCGCTGTCAATTAATACGCGCCGCGAGAGTGAACCTGAC 1200
Db      1328 CAGATCGCTGATTAACAGCGCGCTGTCAATTAATACGCGCCGCGAGAGTGAACCTGAC 1387
Qy      1201 CTGGCAGGATGTGAAGACGAAATGCGAGGGCATGTGTGTTGGCAGGTTTACGCA 1260
Db      1388 CTGGCAGGATGTGAAGACGAAATGCGAGGGCATGTGTGTTGGCAGGTTTACGCA 1447
Qy      1261 ATCGGATTAAGACCGGATACCGGCGTGGCAGTTTGAAGAT 1300
Db      1448 ATCGGATTAAGACCGGATACCGGCGTGGCAGTTTGAAGAT 1487

RESULT 15
LOCUS   AR636191                      1901 bp    DNA          linear    PAT 20-APR-2005
DEFINITION   Sequence 9 from patent US 6855365.
ACCESSION   AR636191
VERSION     AR636191.1  GI:62768874
KEYWORDS
SOURCE      .
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 1901)
AUTHORS    Short,J.M., Kretz,K.A., Gray,K.A., Barton,N.R., Garrett,J.B.,
            O'Donoghue,B. and Mathur,B.J.
TITLE      Recombinant bacterial phytases and uses thereof
JOURNAL    Patent: US 6855365-A 9 15-FEB-2005;
            Diversa Corporation; San Diego, CA
FEATURES
SOURCE     1..1901
```

```
ORIGIN
Query Match      97.3%; Score 1286.2; DB 2; Length 1901;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1292; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

/organism="unknown"
/mol_type="genomic DNA"

1  ATGAAAGCATCTTAATCCCATTTTATCTCTTGTGATCCGTTAACCCCGCATCTGCA 60
188  ATGAAAGCATCTTAATCCCATTTTATCTCTTGTGATCCGTTAACCCCGCATCTGCA 247
61  TTCCTCAGAGTGAAGCCGAGACTGAAGCTGAAGTGTGATGTCACTGTCATGCT 120
248  TTCCTCAGAGTGAAGCCGAGACTGAAGTGTGATGTGATGTGATGTGATGTGATGCT 307
121  GTGCGTCTCCACCAAGGCCAACGCAACTGATGCAAGATGTCACTCCAGACGATGCGCA 180
308  GTGCGTCTCCACCAAGGCCAACGCAACTGATGCAAGATGTCACTCCAGACGATGCGCA 367
181  ACCTGGCCGGTAAACTGGGTGGCTGACACCGGNGGTGGTGAAGTTATGCGCTATCTC 240
368  ACTGGCCGGTAAACTGGGTGGCTGACACCGGNGGTGGTGAAGTTATGCGCTATCTC 427
241  GAGCATTTACCAACGCGCAGCGCTCTGTAGCCGATTTGCTGCGGAAAGGCGTGC 300
428  GAGCATTTACCAACGCGCAGCGCTCTGTAGCCGATTTGCTGCGGAAAGGCGTGC 487
301  CAGTCTGTCAAGTTCGCGATTAATGCTGATGTGCAAGAGCTTACCGTTAAACGCGCA 360
488  CAGTCTGTCAAGTTCGCGATTAATGCTGATGTGCAAGAGCTTACCGTTAAACGCGCA 547
361  GCTTTCGCGCGCGGCTGGCACCTGACTGTGCAATTAACCGTACATCCACGAGGATACG 420
548  GCTTTCGCGCGCGGCTGGCACCTGACTGTGCAATTAACCGTACATCCACGAGGATACG 607
421  TCCAGTCCCGATCCGTTATTTAATCTCTAATAAACTGCGGTTTCCCACTGATACGCG 480
608  TCCAGTCCCGATCCGTTATTTAATCTCTAATAAACTGCGGTTTCCCACTGATACGCG 667
481  AACGTGACTGACGCGATCTCAAGCAAGGCGAGAGGCTCAATTGCTGACTTTACGGGCA 540
668  AACGTGACTGACGCGATCTCAAGCAAGGCGAGAGGCTCAATTGCTGACTTTACGGGCA 727
541  CGGCAACGGCGTTTCGGGAACGGAACGCGGTGCTTAATTTCCGCAATCAACTTGTGC 600
728  CGGCAACGGCGTTTCGGGAACGGAACGCGGTGCTTAATTTCCGCAATCAACTTGTGC 787
601  CTTAAACGTGAGAAACAGACGAAAGCTGTTCAATTAACGAGGATTAACATCGGAATC 660
788  CTTAAACGTGAGAAACAGACGAAAGCTGTTCAATTAACGAGGATTAACATCGGAATC 847
Qy      661 AAGTGAAGCGCCGCAATGTCTCATTTAACGGTGGGTGAAGCTTCATCAATGTCTGACG 720
Db      848 AAGTGAAGCGCCGCAATGTCTCATTTAACGGTGGGTGAAGCTTCATCAATGTCTGACG 907
Qy      721 GAGATATTTCTCTCTGCAACAGCAAGGGAATTCGGAGCCGGGGTGGGGAAGATCAAC 780
Db      908 GAGATATTTCTCTCTGCAACAGCAAGGGAATTCGGAGCCGGGGTGGGGAAGATCAAC 967
Qy      781 GATTCAACACAGTGAACACCTTGCTAGTTTGCAATACGGCAATTTTATTTTGCTACAA 840
Db      968 GATTCAACACAGTGAACACCTTGCTAGTTTGCAATACGGCAATTTTATTTTGCTACAA 1027
Qy      841 CGCAGCGCAGAGTTGCCCGCAGCCGCGCAACCCCGTTATTGATTTGATCAATGCGACG 900
Db      1028 CGCAGCGCAGAGTTGCCCGCAGCCGCGCAACCCCGTTATTGATTTGATCAATGCGACG 1087
Qy      901 TTGAGCGCCCATCCACCGGCAAAAACAGCGCTATGCTGATGACATTAACCACTTCAGTACTG 960
Db      1088 TTGAGCGCCCATCCACCGGCAAAAACAGCGCTATGCTGATGACATTAACCACTTCAGTACTG 1147
961  TTTATTGCGCGACACAGTACTAATCTGGCAATCTCGCGCGCACTGAGCTCAACTGG 1020
```

Db 1148 TTTATCGCCGGACACGATACTAATCTGGCAAACTCGGCGGCGCACTGGAGCTCAACTGG 1207
Qy 1021 ACCCTTCCCGGTCAAGCCGATTAACAGCCGCAAGTGTGAACCTGGTTGAACGCTGG 1080
Db 1208 ACCCTTCCCGGTCAAGCCGATTAACAGCCGCAAGTGTGAACCTGGTTGAACGCTGG 1267
Qy 1081 CGTCGGCTAAGCGATTAACAGCCGATTAACAGCTTGGTGTCTTCCAGACTTTACAG 1140
Db 1268 CGTCGGCTAAGCGATTAACAGCCGATTAACAGCTTGGTGTCTTCCAGACTTTACAG 1327
Qy 1141 CAGATCGTGATTAACAGCCGCTGTCAATTAATACGCGCCGAGAGGTGAACCTGACC 1200
Db 1328 CAGATCGTGATTAACAGCCGCTGTCAATTAATACGCGCCGAGAGGTGAACCTGACC 1387
Qy 1201 CTGGCAGATGTGAAGAGCGAAATGCGCAGGSCATGTGTTGTTGCGAGGTTTACGCA 1260
Db 1388 CTGGCAGATGTGAAGAGCGAAATGCGCAGGSCATGTGTTGTTGCGAGGTTTACGCA 1447
Qy 1261 ATCGTAATGAAGCAGCATACCGGCGTGCACTTGAAT 1300
Db 1448 ATCGTAATGAAGCAGCATACCGGCGTGCACTTGAAT 1487

Search completed: June 14, 2006, 14:10:08
Job time : 7402.95 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2006, 10:01:41 ; Search time 6946.38 Seconds
(without alignments) 10650.337 Million cell updates/sec

Title: US-10-601-319-1
Perfect score: 1322
Sequence: 1 atgaagacgacttaatccc.....atcacatcacatcactca 1323

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 2795965780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est3:*
3: gb_est4:*
4: gb_est5:*
5: gb_est6:*
6: gb_est7:*
7: gb_est8:*
8: gb_est9:*
9: gb_est10:*
10: gb_est11:*
11: gb_est12:*
12: gb_est13:*
13: gb_est14:*
14: gb_est15:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	614.4	46.5	616	2	BJ074127
2	589.8	44.6	853	13	CL662734 PRI0142b
3	528.2	40.0	746	4	CA093060 SCCCL200
4	521	39.4	529	3	BU713770 SUABUG02
5	481.8	36.4	595	5	CF326092 JMT1--05-
6	476.8	36.1	895	8	CN762997 ID0AAASDC
7	438.8	33.2	868	8	CN754382 ID0AAASDC
8	434.8	32.9	864	8	CN757565 ID0AAASDC
9	386.2	29.2	393	7	BE520240 M11B12STM
10	383	29.0	383	7	AM036132 EST274508
11	383	29.0	383	7	AM036134 EST274510
12	348.6	26.4	531	3	BU713769 SUABUG01
13	304.6	23.0	354	1	AA545747 HBMSP184-
14	295.8	22.4	706	8	CN759004 ID0AAASDC
15	248	18.8	707	2	BU619443 BU619443
16	216.8	16.4	320	7	BE436403 EST407481
17	208	15.7	208	2	BM409016 EST583343
18	208	15.7	208	2	BM412806 EST587133
19	198.8	15.0	546	2	BU040796 BU040796

c	20	169.4	12.8	179	7	BE461872	BE461872
c	21	160.8	12.2	508	2	BU030203	BU030203
c	22	157.6	11.9	657	2	BE457132	BE457132
c	23	156.4	11.8	1171	8	CV672454	CV672454
c	24	143.2	10.8	254	7	BE520241	BE520241
c	25	103.4	7.8	769	8	CO386885	CO386885
c	26	88.8	6.7	863	8	CV672453	CV672453
c	27	57.2	4.3	772	14	DX010757	DX010757
c	28	55.2	4.2	821	13	CL665831	CL665831
c	29	55.2	4.2	842	13	CL665745	CL665745
c	30	49.8	3.8	470	9	DN201839	DN201839
c	31	48.8	3.7	1175	9	CX943192	CX943192
c	32	43.8	3.3	392	10	H60613	H60613
c	33	41.2	3.1	733	2	BT152280	BT152280
c	34	39.8	3.0	346	4	BY053637	BY053637
c	35	39.8	3.0	350	4	BY041909	BY041909
c	36	39.8	3.0	365	4	BY056036	BY056036
c	37	39.8	3.0	389	4	CA566872	CA566872
c	38	39.8	3.0	431	1	AT553154	AT553154
c	39	39.8	3.0	539	4	BX517075	BX517075
c	40	39.8	3.0	602	7	BB610221	BB610221
c	41	39.8	3.0	607	1	AA521521	AA521521
c	42	39.8	3.0	632	4	BY738708	BY738708
c	43	39.8	3.0	646	4	BY737657	BY737657
c	44	39.8	3.0	2863	6	AK146141	AK146141
c	45	39.8	3.0	2863	6	AK168364	AK168364

ALIGNMENTS

RESULT 1
BU074127/c 616 bp mRNA linear EST 29-SEP-2003
LOCUS BU074127 NIBB Mochii normalized Xenopus tailbud library Xenopus
DEFINITION laevis cDNA clone X1090106 5', mRNA sequence.
ACCESSION BU074127 GI:17504316
VERSION BU074127.1 GI:17504316
KEYWORDS
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus; Xenopus.
1 (bases 1 to 616)
Kiteyama, A., Terashima, C., Mochii, M., Ueno, N., Shin-I, T. and
Kohara, Y.
Expressed genes in X. laevis embryo
Unpublished (2001)
Contact: Tadashi Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tsun@genes.nig.ac.jp
The information of this clone is available through the following
URL.
http://xenopus.nibb.ac.jp.

FEATURES
source location/Qualifiers
1..616
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="X1090106"
/issue_type="whole embryo"
/dev_stage="stage 25"
/clone_lib="NIBB Mochii normalized Xenopus tailbud
library"

ORIGIN

Query Match 46.5%; Score 614.4; DB 2; Length 616;
Best Local Similarity 99.8%; Pred. No. 8.3e-180;

REFERENCE 1 (bases 1 to 746)
AUTHORS Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
TITLE The libraries that made SUCEST
JOURNAL Genet. Mol. Biol. 24 (1-6), 1-7 (2001)
COMMENT Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: pattud@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 001 row: D column: 10
Seq primer: M13/Forward primer.
Location/Qualifiers

FEATURES
source 1..746
/organism="Saccharum officinarum"
/mol_type="rRNA"
/db_xref="taxon:4547"
/clone="SCCCL2001D10"
/lab_host="XLIBlue MRP"
/clone_1lb="CL2"
/note="Organ: Pool of sugarcane calli submitted to low
(40C) and high (37 C) temperature stress; Vector:
pBluscript; Site 1: EcoRI; Site 2: XhoI; An
unidirectional cDNA library generated from [pool of
sugarcane calli submitted to low (40C) and high (37 C)
temperature stress]. cDNA was prepared from poly(A+ mRNA
using ZAP - cDNA Synthesis Kit (Stratagene). The
double-strand cDNAs were fractionated in a sepharose CL-2B
40cm-column and fragments sizing between 0.8 and 1.5 Kb
were directionally cloned into the vector. Details of
each source of RNA and library construction can be
obtained at http://sucest.fad.ic.unicamp.br/public"

ORIGIN

Query Match 40.0%; Score 528.2; DB 4; Length 746;
Best Local Similarity 97.0%; Pred. No. 7.1e-153;
Matches 547; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

OY 595 TTGTCCTTAAACGAGAAACAGGACGAACTGTCTTAAACGAGCATTAACCATG 654
DB 184 TTGCTACTGTGAGGAAACAGGACGAACTGTCTTAAACGAGCATTAACCATG 243
OY 655 GAACTCAAGGTGAGCGCGACATGCTCATTAACCGGTGCGTAAGCTCGCATCAATG 714
DB 244 GAACTCAAGGTGAGCGCGACATGCTCATTAACCGGTGCGTAAGCTCGCATCAATG 303
OY 715 CTGAGCGAGATTTTCTCTGCAACAGCAAGGAATCCGAGCCGGGTGGGGAAG 774
DB 304 CTGAGCGAGATTTTCTCTGCAACAGCAAGGAATCCGAGCCGGGTGGGGAAG 363
OY 775 ATACCGATTACACAGTGAACCTTGCAATTTGCAATGCGCAATTTATTTG 834
DB 364 ATACCGATTACACAGTGAACCTTGCAATTTGCAATGCGCAATTTATTTG 423
OY 835 CTACACGACGCGCAGAGGTTGCGCGACGCGCAACCGCTTATGATTTGATCATG 894
DB 424 CTACACGACGCGCAGAGGTTGCGCGACGCGCAACCGCTTATGATTTGATCATG 483
OY 895 GCAAGCTTGAGCGCCCATTCACCGCAAAAACGCGGTATGTGTGACATTACCACTTCA 954
DB 484 ACAGGTTGAGCGCCCATTCACCGCAAAAACGCGGTATGTGTGACATTACCACTTCA 543
OY 955 GTAAGTTATTTGCGGACAGATCTAATCTGGCAATCTGGGCGGCACTGAGAGTTC 1014
DB 544 GTGCTGTTATTCGCGGACAGATCTAATCTGGCAATCTGGGCGGCACTGAGAGTTC 603
OY 1015 AACTGAGCGCTTCCGCTCAGCCGATTAACGCGCGGAGGTGTGAATGGTGTGAA 1074

DB 604 AACTGAGCGCTTCCGCTCAGCCGATTAACGCGCGGAGGTGTGAATGGG-TTGAA 662
OY 1075 CGCTGCGCTGCGCTAAGCATTAACGCGCATGATTTCCGTTCCAGACT 1134
DB 663 CGCTGCGCTGCGCTAAGCATTAACGCGCATGATTTCCGTTCCAGACT 722
OY 1135 TTACGACAGATCGTGATTAACG 1158
DB 723 TTACGACAGATCGTGATTAACG 746

RESULT 4
BU713770 529 bp mRNA linear EST 23-OCT-2003
LOCUS BU713770.1 GI:28321126
DEFINITION SUGARB02 Adult SJC 7/94 Schistosoma japonicum cDNA similar to
pDB1DXP/A Chain A, Crystal Structure Of Phytate Complex Of
Escherichia Coli Phytase At Ph 6.6. Phytate Is Bound With Its
3-Phosphate In The Active Site. Hg2+ Cation Acts As An
Intermolecular Bridge, mRNA sequence.

ACCESSION BU713770.1 GI:28321126
VERSION BU713770.1
KEYWORDS EST.
SOURCE Schistosoma japonicum
ORGANISM Schistosoma japonicum
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
Strigeidida; Schistosomatidae; Schistosomatidae; Schistosoma.

REFERENCE

1 (bases 1 to 529)
Hu, W., Yan, Q., Shen, D. K., Liu, F., Zhu, Z. D., Song, H. D., Xu, X. R.,
Wang, Z. J., Rong, Y. P., Zeng, L. C., Wu, D., Zhang, X., Wang, J. J.,
Xu, X. N., Wang, S. Y., Fu, G., Zhang, X. L., Wang, Z. Q., Birdley, P. J.,
McManus, D. P., Xue, C. L., Feng, Z., Chen, Z. and Han, Z. G.
Evolutionary and biomedical implications of a Schistosoma japonicum
complementary DNA resource
Nat. Genet. 35 (2), 139-147 (2003)
12973349

JOURNAL

PUBMED

COMMENT

Contact: Zenguan Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex. 45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn.

FEATURES
source 1..529
Location/Qualifiers

/organism="Schistosoma japonicum"
/mol_type="rRNA"
/strain="Chinese (Anhui) strain"
/db_xref="taxon:6182"
/sex="Male and female"
/tissue_type="Whole body"
/dev_stage="Adult worms"
/lab_host="Mouse and rabbit"
/clone_1lb="Adult SJC 7/94"
/note="Vector: Lambda ZAP-II XR.; Site 1: EcoR I; Site 2:
XhoI I; Several hundred adult Schistosoma japonicum
(Anhui, P. R. China, strain), of mixed sex, were perfused
from the mesenteries of experimentally infected mice and
rabbits at the Queensland Institute of Medical Research,
Brisbane, Australia (QIMR), and stored for several months
in liquid nitrogen. Subsequently, mRNA was isolated at the
QIMR from lysates of these worms by oligo dt
chromatography, using a kit from Pharmacia. The mRNA was
then shipped to Clontech, Palo Alto, CA, USA, who
constructed a cDNA library. First strand synthesis was
primed with an oligo-dt-XhoI primer and synthesized using
M-MuV reverse transcriptase. Second strand synthesis was
accomplished with RNase H and T4 DNA polymerase. The
double stranded cDNA was ligated to EcoRI linkers,
digested with EcoRI and XhoI, and ligated into the
phagemid vector lambda ZAP II XR. After construction of
this directional library by Clontech, it was returned to
the QIMR. During analysis of the library at the QIMR, we
have found that a small percentage, 2% to 3%, of the

clones contain inserts that appear to be highly homologous to sequences from salmonid fishes, as determined by homology comparisons using BLAST and by Southern hybridization analysis to genomic DNA from salmon (Sigma Chemical Co., St. Louis, MO) under stringent washing conditions. The remainder of the clones appear to contain *S. japonicum* sequences."

ORIGIN

Query Match 39.4%; Score 521; DB 3; Length 529;
Best Local Similarity 99.1%; Pred. No. 1.2e-150;
Matches 524; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 670 GCCGCAATGCTCTATTAAACCGGTGCGGTAAAGCTCGCATATGCTGACGAGATATTT 729
Db 1 GCCGCAATGCTCTATTAAACCGGTGCGGTAAAGCTCGCATATGCTGACGAGATATTT 60
Oy 730 CTCCTGCAACAAGCAAGAGGAATGCCGAGCCGGGGTGGGAAGATCAACGATTCAAC 789
Db 61 CTCCTGCAACAAGCAAGAGGAATGCCGAGCCGGGGTGGGAAGATCAACGATTCAAC 120
Oy 790 CAGTGGAAACACCTTCTAATGTTGATTAACGCGCAATTTTATTGCTACAAAGCAGCC 849
Db 121 CAGTGGAAACACCTTCTAATGTTGATTAACGCGCAATTTTATTGCTACAAAGCAGCC 180
Oy 850 GAGGTTCGCCGAGCCGCGCCACCCGTTATTTGATTCATGATGAGCGCTGACGCGC 909
Db 181 GAGGTTCGCCGAGCCGCGCCACCCGTTATTTGATTCATGATGAGCGCTGACGCGC 240
Oy 910 CATCCACCGCAAAAACAGCGGTATGATGATTAACCACTTCAGTACTGTTTATTCG 969
Db 241 CATCCACCGCAAAAACAGCGGTATGATGATTAACCACTTCAGTACTGTTTATTCG 300
Oy 970 GGAACCAATCTAATCTGCAAAATTCGCGCGCGCACTGAGACTCAATCGACGTTTCC 1029
Db 301 GGAACCAATCTAATCTGCAAAATTCGCGCGCGCACTGAGACTCAATCGACGTTTCC 360
Oy 1030 GGTGACCGGATTAACAGCCCGCAGGTGTAATCTGTTTGAACGCGCGGTGCGCTA 1089
Db 361 GGTGACCGGATTAACAGCCCGCAGGTGTAATCTGTTTGAACGCGCGGTGCGCTA 420
Oy 1090 AGCGATTAACAGCGATGATTCAGGTTTCGTGTCCTTCAGACTTTACAGCAATGCGT 1149
Db 421 AGCGATTAACAGCGATGATTCAGGTTTCGTGTCCTTCAGACTTTACAGCAATGCGT 480
Oy 1150 GATTAACAGCGCTGCTATTAAATCGCGCCCGGAGAGGTAACTGA 1198
Db 481 GATTAACAGCGCTGCTATTAAATCGCGCCCGGAGAGGTAACTGA 529

RESULT 5
CP326092/c 595 bp mRNA linear EST 18-AUG-2003
LOCUS JMT1--05-B09.g1 AtJMT-overexpressing transgenic rice lambda phage
DEFINITION cDNA library (JMT1) Oryza sativa (japonica cultivar-group) cDNA
clone JMT1--05-B09, mRNA sequence.

ACCESSION CP326092
VERSION CP326092.1 GI:33800445

KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)

REFERENCE 1 (bases 1 to 595)
Kim, J.-S., Jun, K.-M., Cheong, P.-J., Kim, M.-J., Lee, T.-H., Shin, Y.-C.,
Song, S.-I., Kim, J.-K., Kim, Y.-K. and Nahm, B.-H.
Large-scale sequencing analysis of rice ESTs

Unpublished (2003)
Contact: Nahm B.H.
Genomics and Genetics Institute, GreenGene Biotech Inc., Division
of Bioscience and Bioinformatics, Myongji University
Yongin, Kyonggi, Korea

Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.
Location/Qualifiers
1..595
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:39947"
/clone="JMT1--05-B09"
/tissue_type="leaf"
/dev_stage="14 days after germination"
/lab_host="E. coli SOLR"
/clone_id="AtJMT-overexpressing transgenic rice lambda
phage cDNA library (JMT1)"
/note="Vector: pBluescript SK(+); Site 1: EcoRI; Site 2:
XhoI; cDNA was inserted into lambda Uni-ZAP XR vector at 5'
end with EcoRI and 3' end with XhoI site. mRNA was
prepared from Arabidopsis Jasmonate Carbonyl
methyltransferase overexpression line."

FEATURES

SOURCE

Query Match 36.4%; Score 481.8; DB 5; Length 595;
Best Local Similarity 98.6%; Pred. No. 2.1e-138;
Matches 486; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

ORIGIN

Oy 808 AGTTGCAATACGGCAATTTTATTGCTACAAAGCAAGCAGAGTTGCCGAGCCG 867
Db 595 AGTTGCAATACGGCAATTTTATTGCTACAAAGCAAGCAGAGTTGCCGAGCCG 536
Oy 868 GCCACCCCGTATTGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 927
Db 535 GCCACCCCGTATTGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 476
Oy 928 GCGTATGCTGATTAACCACTTCAGTACTGTTTATGCGCGGACAGATCAATCTG 987
Db 475 GCGTATGCTGATTAACCACTTCAGTACTGTTTATGCGCGGACAGATCAATCTG 416
Oy 988 GCAATCTCGCGCGGCGGCACTGAGCTCACTGAGCGCTCCCGGTACGCGGATTAAC 1047
Db 415 GCAATCTCGCGCGGCGGCACTGAGCTCACTGAGCGCTCCCGGTACGCGGATTAAC 356
Oy 1048 CCGCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1107
Db 355 CCGCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 296
Oy 1108 ATTGAGGTTTGGCTGCTTTTCAAGCTTTAAGAGATGCTGATTAACCGCCCTG 1167
Db 295 ATTGAGGTTTGGCTGCTTTTCAAGCTTTAAGAGATGCTGATTAACCGCCCTG 236
Oy 1168 TTAATACGCGCGCGGAGAGAGTGAACCTGCGAGATGATGAAGAGCAATGCG 1227
Db 235 TTAATACGCGCGCGGAGAGAGTGAACCTGCGAGATGATGAAGAGCAATGCG 176
Oy 1228 CAGGCAATGCTGCTGTTGAGAGGTTTAAACGAAATCGTAATGAAGCAGATAC 1287
Db 175 CAGGCAATGCTGCTGTTGAGAGGTTTAAACGAAATCGTAATGAAGCAGATAC 116
Oy 1288 TGCACTTTGAGAT 1300
Db 115 TGCACTTTGAT 103

RESULT 6

CN762997/c 895 bp mRNA linear EST 20-MAY-2004
LOCUS ID0AAA5DC09RM1 ApMS Acyrthosiphon pisum cDNA clone ID0AAA5DC09 5',
DEFINITION mRNA sequence.

ACCESSION CN762997
VERSION CN762997.1 GI:47536920

KEYWORDS EST.
SOURCE Acyrthosiphon pisum (pea aphid)
ORGANISM Acyrthosiphon pisum

REFERENCE

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes; Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.

1 (bases 1 to 895)
Hunter, W., Martinez-Torres, D., Rahe, Y., Sabater-Munoz, B., Stern, D., Tagu, D. and Wincker, P.

TITLE

An expressed sequence tags database for the pea aphid *Acyrthosiphon pisum*

JOURNAL

Unpublished (2004)

Contact: D. Tagu

COMMENT

INRA Rennes
UMR Bio3P, BP 35327, F-35653 Le Rheu Cedex France

Tel: +33.2.23.48.51.65

Fax: +33.2.23.48.51.50

Risk of contamination by bacterial sequences from obligatory (Buchnera) or facultative endosymbionts. These sequences were obtained in the frame of the International Consortium of Aphid Genomics in collaboration with Genoscope

PCR Primers

FORWARD: CAGGAACAGCATGACG

Plate: 5 row: C column: 9.

FEATURES

Location/Qualifiers

1..895

/organism="Acyrthosiphon pisum"

/mol_type="mRNA"

/cultivar="developmentstage"

/db_xref="taxon:7029"

/clone="ID0AA12DE01"

/tissue_type="whole insect"

/dev_stage="nymphs and adults (parthenogenetic females)"

/lab_host="X11-Blue"

/clone_1ib="ApMS"

/note="Vector: pBS-SK minus; Site 1: EcoRI; Site 2: XhoI; Sample name: ID0AA1; Plant growth place: Department of Ecology & Evolutionary Biology, Princeton University; Soil conditions: Soil; Sowing date: 01/06/1999; Harvesting date: 01/06/1999; Stress date: no stress; Description: Aphid inoculated on one-week old *Vicia faba* under non-sterile conditions. All parthenogenetic stages and both winged and wingless adults were collected for library construction; experimental condition: long photoperiod (16-hr light/8-hr dark at 18 C)"

ORIGIN

Query Match 36.1%; Score 476.8; DB 8; Length 895;

Best Local Similarity 98.6%; Pred. No. 8.4e-137;

Matches 481; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

813 GCATAACGGCAATTTATTTGCTACACGACGCGGATGCGCGCGCCAC 872
895 GCATAACGGCAATTTATTTGCTACACGACGCGGATGCGCGCGCCAC 836
873 CCCGTTATTTGATGATGATGATGATGATGATGATGATGATGATGAT 932
835 CCCGTTATTTGATGATGATGATGATGATGATGATGATGATGATGAT 776
933 TGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 992
775 TGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 716
993 TCTCGGCGGCGCATGAGCTGACGAGCTTCCCGGTGACGCGGATTAAC 1052
715 TCTCGGCGGCGCATGAGCTGACGAGCTTCCCGGTGACGCGGATTAAC 656
1053 AGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1112
655 AGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 596
1113 GGTTCGCTGCTGCTTCCACATTTACAGAGATGCTGATTAACGCGCTG 1172
595 GGTTCGCTGCTGCTTCCACATTTACAGAGATGCTGATTAACGCGCTG 536
1173 TACGCGCGCGGAGAGTGAACCTGCGAGAGATGTAAGAGCAAAATGCG 1232

||||| 535 TACGCGCGCGGAGAGTGAACCTGCGAGAGATGTAAGAGCAAAATGCGAGG 476

QY 1233 CATGCTGCTGTTGGCAGGTTTACGCAATCGTAATGAAGACGATACCGGCTGCAG 1292

DB 475 CATGCTGCTGTTGGCAGGTTTACGCAATCGTAATGAAGACGATACCGGCTGCAG 416

QY 1293 TTTGAGAT 1300

DB 415 TTTGTAAT 408

RESULT 7

CN754382/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

INRA Rennes

UMR Bio3P, BP 35327, F-35653 Le Rheu Cedex France

Tel: +33.2.23.48.51.65

Fax: +33.2.23.48.51.50

Risk of contamination by bacterial sequences from obligatory (Buchnera) or facultative endosymbionts. These sequences were obtained in the frame of the International Consortium of Aphid Genomics in collaboration with Genoscope

PCR Primers

FORWARD: CAGGAACAGCATGACG

Plate: 12 row: B column: 1.

FEATURES

Location/Qualifiers

1..868

/organism="Acyrthosiphon pisum"

/mol_type="mRNA"

/cultivar="developmentstage"

/db_xref="taxon:7029"

/clone="ID0AA12DE01"

/tissue_type="whole insect"

/dev_stage="nymphs and adults (parthenogenetic females)"

/lab_host="X11-Blue"

/clone_1ib="ApMS"

/note="Vector: pBS-SK minus; Site 1: EcoRI; Site 2: XhoI; Sample name: ID0AA1; Plant growth place: Department of Ecology & Evolutionary Biology, Princeton University; Soil conditions: Soil; Sowing date: 01/06/1999; Harvesting date: 01/06/1999; Stress date: no stress; Description: Aphid inoculated on one-week old *Vicia faba* under non-sterile conditions. All parthenogenetic stages and both winged and wingless adults were collected for library construction; experimental condition: long photoperiod (16-hr light/8-hr dark at 18 C)"

ORIGIN

Query Match 33.2%; Score 438.8; DB 8; Length 868;

Best Local Similarity 98.3%; Pred. No. 6.1e-125;

Matches 454; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 839 AACGCAACCCAGAGTTTCCCGCAGCGCCACCCGTTATTGATTGATGACGAG 898
868 AACGCAACCCAGAG-TTCCCGCAGCGCCACCCCGTTATTGATTGATGACGAG 810

QY 899 CGTTGAGCCCGCATCCACCGGAAAAACAGGGGTATGTTGATCAATTACCACTTCAGAC 958
DB 809 CGTTGAGCCCGCATCCACCGGAAAAACAGGGGTATGTTGATCAATTACCACTTCAGAC 750
QY 959 TGTATATGCGCGACACGATATCTATCTGCAAAATCTCGCGCGCACTGAGCTCACT 1018
DB 749 TGTATATGCGCGACACGATATCTATCTGCAAAATCTCGCGCGCACTGAGCTCACT 690
QY 1019 GAGCGCTTCCGGGTGACCGCGATTAACAGCCGCCAGGTGTGAACCTGTTTGAACGCT 1078
DB 689 GAGCGCTTCCGGGTGACCGCGATTAACAGCCGCCAGGTGTGAACCTGTTTGAACGCT 630
QY 1079 GAGCGCTTCCGGGTGACCGCGATTAACAGCCGCCAGGTGTGAACCTGTTTGAACGCT 1138
DB 629 GAGCGCTTCCGGGTGACCGCGATTAACAGCCGCCAGGTGTGAACCTGTTTGAACGCT 570
QY 1139 AGCAGATCGTATTAACAGCGCGCTGTCAATTAAATACCGCCCGGAGAGGTGAACCTGA 1198
DB 569 AGCAGATCGTATTAACAGCGCGCTGTCAATTAAATACCGCCCGGAGAGGTGAACCTGA 510
QY 1199 CCTGCGCAGATGTGAAGAGCGAAATGCGCAGGCGATGTCTTCTTGCGAGGTTTACGC 1258
DB 509 CCTGCGCAGATGTGAAGAGCGAAATGCGCAGGCGATGTCTTCTTGCGAGGTTTACGC 450
QY 1259 AAATGTGAATGAAGACGACGATACCGGCGGTGAGTTGAGAT 1300
DB 449 AAATGTGAATGAAGACGACGATACCGGCGGTGAGTTGAGAT 408

RESULT 8
LOCUS CN757565/c 864 bp mRNA linear EST 20-MAY-2004
DEFINITION ID0AAALDf09RM1 ApMS Acyrthosiphon pisum cDNA clone ID0AAALDf09 5',
mRNA sequence.

ACCESSION CN757565
VERSION CN757565.1 GI:47531488
KEYWORDS EST.

SOURCE Acyrthosiphon pisum (pea aphid)
ORGANISM Acyrthosiphon pisum
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.

REFERENCE 1 (bases 1 to 864)
Hunters, W., Martinez-Torres, D., Rabhe, Y., Sabater-Munoz, B.,
Stern, D., Tagu, D. and Wincker, P.
An expressed sequence tags database for the pea aphid Acyrthosiphon
pisum

TITLE Unpublished (2004)
JOURNAL Contact: D. Tagu
COMMENT INRA Rennes
UMR BIO3P, BP 35327, F-35653 Le Rheu Cedex France
Tel: +33.2.23.48.51.65
Fax: +33.2.23.48.51.50

FEATURES
source
1. 864
/organism="Acyrthosiphon pisum"
/mol_type="mRNA"
/catalvar="developmentstage"
/db_xref="taxon:7029"
/clone="ID0AAALDf09"
/tissue_type="whole insect"
/dev_stage="nymphs and adults (parthenogenetic females)"
/lab_host="XL1-Blue"
/clone_lib="ApMS"
/note="Vector: pBS-SK minus; site_1: EcoRI; site_2: XhoI;

Sample name: ID0AAA ; Plant growth place: Department of
Ecology & Evolutionary Biology, Princeton University ;
Soil conditions: Soil ; Sowing date: 01/06/1999 ;
Harvesting date: 01/06/1999 ; Stress date: no stress ;
Description: Aphids inoculated on one-week old *Vicia faba*
under non-sterile conditions. All parthenogenetic stages
and both winged and wingless adults were collected for
library construction. ; experimental condition: long
photoperiod (16-hr light/8-hr dark at 18 °C)

Query Match 32.9%; Score 434.8; DB 8; Length 864;
Best Local Similarity 98.3%; Pred. No. 1,1e-123;
Matches 450; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 843 CAGCCGAGAGTTGCCCGGACCGCCGACCCCTTTATGGATTGATTCATGAGAGGTT 902
DB 864 CAGCCGAGAGTTGCCCGGACCGCCGACCCCTTTATGGA-TTGATCAAGACAGCGTT 806
QY 903 GAGCCCGCATCCACCGGAAAAACAGGGGTATGTTGATCAATTACCACTTCAGTCTGT 962
DB 805 GAGCCCGCATCCACCGGAAAAACAGGGGTATGTTGATCAATTACCACTTCAGTCTGT 746
QY 963 TATGCGGACACGATATCTAATCTGCAAAATCTGCGCGGCACTGAGACTCACTGAC 1022
DB 745 TATGCGGACACGATATCTAATCTGCAAAATCTGCGCGGCACTGAGACTCACTGAC 686
QY 1023 GCTTCCCGGTATGCGCGATTAACAGCCGCCAGGTGTGAACCTGTTTGAACCTGCGG 1082
DB 685 GCTTCCCGGTATGCGCGATTAACAGCCGCCAGGTGTGAACCTGTTTGAACCTGCGG 626
QY 1083 TGGGCTAAGGATTAACAGCCGATGATTCAGTTTCTGCTGCTTCCAGACTTACAGCA 1142
DB 625 TGGGCTAAGGATTAACAGCCGATGATTCAGTTTCTGCTGCTTCCAGACTTACAGCA 566
QY 1143 GATGCGTATTAACAGCCGCTGTCAATTAAATACCGCCCGGAGAGGTGAACCTGACCT 1202
DB 565 GATGCGTATTAACAGCCGCTGTCAATTAAATACCGCCCGGAGAGGTGAACCTGACCT 506
QY 1203 GCGCAGATGTGAAGAGCGAAATGCGCAGGCGATGTCTTCTTGCGAGGTTTACGCAAT 1262
DB 505 GCGCAGATGTGAAGAGCGAAATGCGCAGGCGATGTCTTCTTGCGAGGTTTACGCAAT 446
QY 1263 CGTGAATGAAGACGACGATACCGGCGGTGAGTTGAGAT 1300
DB 445 CGTGAATGAAGACGACGATACCGGCGGTGAGTTGAGAT 408

RESULT 9
LOCUS BE520240/c 393 bp mRNA linear EST 19-MAR-2001
DEFINITION M1B158TM Arabidopsis developing seed Arabidopsis thaliana cDNA
clone M1B12 5', mRNA sequence.

ACCESSION BE520240
VERSION BE520240.1 GI:978242
KEYWORDS EST.

SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 393)
White, J.A., Todd, J., Newman, T., Focks, N., Girke, T., Martinez de
Ibarra, O., Javoriski, J.G., Ohlrogge, J. and Benning, C.
A new set of Arabidopsis expressed sequence tags from developing
seeds. The metabolic pathway from carbohydrates to seed oil
Plant Physiol. 124 (4), 1582-1594 (2000)

JOURNAL PUBMED
COMMENT Contact: Benning, C
Dept. of Biochemistry & Molecular Biology
Michigan State University
224 Biochemistry, Michigan State University, East Lansing, MI
48824, USA

Tel: 517 355 1609
 Fax: 517 353 9334
 Email: benning@osu.edu
 Michigan State University DNA Sequencing Facility Arabidopsis
 Biological Resource Center, The Ohio State University, 309 Botany &
 Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210 USA, FAX:
 6142920603 TEL: 6142929371.
 Location/Qualifiers

FEATURES

source

1..393
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /ecotype="Columbia"
 /db_xref="taxon:3702"
 /clone="M1B12"
 /tissue_type="seed"
 /dev_stage="5-13 days after flowering"
 /lab_host="E.coli"
 /clone_lib="Arabidopsis developing seed"
 /note="Organ: Developing seed; Vector: pBluescript SK-;
 Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 29.2%; Score 386.2; DB 7; Length 393;
 Best Local Similarity 98.7%; Pred. No. 1.3e-108;
 Matches 388; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

319 ATTATTCGATGTGACGAGGCTACCCGTTAAACAGCGGAGCCCTTCGCCGCCGGGCTG 378
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 393 ATTATTCGATGTGACGAGGCTACCCGTTAAACAGCGGAGCCCTTCGCCGCCGGGCTG 334
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 379 GCACCTGACTGTGCATATACCGTACATACCCAGCGAGATACGTCAGTCCCGATCCGTTA 438
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 333 GCACTGACTGTGCATATACCGTACATACCCAGCGAGATACGTCAGTCCCGATCCGTTA 274
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 439 TTTATTCCTCTTAAAACTGGCGCTTGGCCAACTGGATTAACGCGAACGTCGACGCGATC 498
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 273 TTTATTCCTCTTAAAACTGGCGCTTGGCCAACTGGATTAACGCGAACGTCGACGCGATC 214
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 499 CTCACAGGCGGACGAGGCTCAATTGCTGACTTTACCGGCGATCGGCAACGGCGTTTCCG 558
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 213 CTCACAGGCGGACGAGGCTCAATTGCTGACTTTACCGGCGATCGGCAACGGCGTTTCCG 154
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 559 GAACCTGAAACGGGCTTAAATTTTCGCAATCAAACTTGTGCTTAAACGTAGAAACAG 618
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 153 GAACCTGAAACGGGCTTAAATTTTCGCAATCAAACTTGTGCTTAAACGTAGAAACAG 94
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 619 GACGAAGCTGTCTTAAACGAGGACTTACCATCGAACTCAAGGTAGCGCCGACAT 678
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 93 GACGAAGCTGTCTTAAACGAGGACTTACCATCGAACTCAAGGTAGCGCCGACAT 34
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 679 GTCTCATTAACCGGTGCGGTAAAGCTTCGATCA 711
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 33 GTCTCATTAACCGGTGCGGTAAAGCTTCGATCA 1

RESULT 10 383 bp mRNA linear EST 18-MAY-2001
 LOCUS AM036132
 DEFINITION EST274508 tomato seed, TANU Lycopersicon esculentum cDNA clone
 CLEB1E23 similar to periplasmic phosphoanhydride phosphohydrolase
 precursor, putative, mRNA sequence.
 ACCESSION AM036132
 VERSION AM036132.1 GI:5894811
 KEYWORDS EST.
 SOURCE Lycopersicon esculentum (Solanum lycopersicum)
 ORGANISM Lycopersicon esculentum
 BUKARYOTA; VIRIDIPLANTAE; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamids; Solanales; Solanales; Solanales; Solanum; Lycopersicon.
 1 (bases 1 to 383)
 Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T.,
 Holt,I.E., Liang,F., Upson,J., Rønning,C.M., Craven,M.B.,
 Fujii,C.Y., Bowman,C.L., Niernan,W., Fraser,C.M., Venter,J.C.,

TITLE

JOURNAL

COMMENT

Martin,G.B., Tanksey,S.D. and Giovannoni,J.
 Generation of ESTs from tomato seed tissue
 Unpublished (1999)
 Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 3 prime sequence.
 Location/Qualifiers

FEATURES

source

1..383
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="CLEB1E23"
 /tissue_type="seeds"
 /dev_stage="quiescent seed"
 /lab_host="XLI-Blue MRF"
 /clone_lib="tomato seed, TANU"
 /note="Vector: pBluescript SK(-); Site_1: EcoRI, Site_2:
 XhoI; cDNA - Tomato Seed EST Library. Directionally cloned
 cDNAs inserted into pBluescript SK(-) at 5' end with
 EcoRI and 3' end with XhoI site."

ORIGIN

Query Match 29.0%; Score 383; DB 7; Length 383;
 Best Local Similarity 100.0%; Pred. No. 1.3e-107;
 Matches 383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

348 TAAACAGGCGAAGCTTCGCGCGGGCTGGACCTGATGTGAAATTAACGTTACATAC 407
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 1 TAAACAGGCGAAGCTTCGCGCGGGCTGGACCTGATGTGAAATTAACGTTACATAC 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 408 CCAGCAGATACGTCCAGTCCGATCCGTTATTTATCTCTTAAACCTGCGTTGCCA 467
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 61 CCAGCAGATACGTCCAGTCCGATCCGTTATTTATCTCTTAAACCTGCGTTGCCA 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 468 ACTGATTAACGGAACGTGACTGACGCAATCTTCAGCAGGCGAGGCTCAATTGCTGA 527
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 121 ACTGATTAACGGAACGTGACTGACGCAATCTTCAGCAGGCGAGGCTCAATTGCTGA 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 528 CTTTACCGGCGATCGGCAACGCGCTTTCGGAATCGGAACGGGCTTAAATTTCCGA 587
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 191 CTTTACCGGCGATCGGCAACGCGCTTTCGGAATCGGAACGGGCTTAAATTTCCGA 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 588 ATCAACTTGTGCTTAAACGTGAGAAACAGGACGAAGCTGTTCATTAAACGAGCAAT 647
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 241 ATCAACTTGTGCTTAAACGTGAGAAACAGGACGAAGCTGTTCATTAAACGAGCAAT 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 648 ACCATCGGAACCTCAAGGTGAGCGCGGACCAATGTCTCATTAACCGGTGCGGTAAAGCTTGC 707
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 301 ACCATCGGAACCTCAAGGTGAGCGCGGACCAATGTCTCATTAACCGGTGCGGTAAAGCTTGC 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 708 ATCAATGCTGAACGAGATATTTTC 730
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 361 ATCAATGCTGAACGAGATATTTTC 383

RESULT 11 383 bp mRNA linear EST 18-MAY-2001
 LOCUS AM036134/C
 DEFINITION EST274510 tomato seed, TANU Lycopersicon esculentum cDNA clone
 CLEB1E23 similar to periplasmic phosphoanhydride phosphohydrolase
 precursor, putative, mRNA sequence.
 ACCESSION AM036134
 VERSION AM036134.1 GI:5894813
 KEYWORDS EST.
 SOURCE Lycopersicon esculentum (Solanum lycopersicum)
 ORGANISM Lycopersicon esculentum
 BUKARYOTA; VIRIDIPLANTAE; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamids; Solanales; Solanales; Solanales; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 383)
 AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vialon,T.,
 Holt,I.E., Ilang,F., Upton,J., Roming,C.M., Craven,M.B.,
 Fujii,C.Y., Bowman,C.L., Nierman,M., Fraser,C.M., Venter,J.C.,
 Martin,G.B., Tanksley,S.D. and Giovannoni,J.
 TITLE Generation of ESTs from tomato seed tissue
 JOURNAL Unpublished (1999)
 COMMENT Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.

FEATURES
 source
 1..383
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /culti_var="TA96"
 /db_xref="taxon:4081"
 /clone="CLB1E23"
 /tissue_type="seeds"
 /dev_stage="quiescent seed"
 /lab_host="XLI-Blue MRP"
 /clone_id="tomato seed, TAMU"
 /note="Vector: pBluescript SK(-); Site 1: EcoRI, Site 2:
 XhoI; cDNA - Tomato Seed EST Library, Directionally cloned
 cDNAs inserted into pBluescript SK(-) at 5' end with
 EcoRI and 3' end with XhoI site."

ORIGIN
 Query Match 29.0%; Score 383; DB 7; Length 383;
 Best Local Similarity 100.0%; Pred. No. 1.3e-107;
 Matches 383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 348 TAAACAGCGGAGGCTTCGCCGCGGCGTGGCACTGTCGATTAACCGTACATAC 407
 Db TAAACAGCGGAGGCTTCGCCGCGGCGTGGCACTGTCGATTAACCGTACATAC 324
 OY 408 CCAGGCAATACGTGCTGATCCGATCCGTTATTTAATCCTCTAAACCTGGGTTTGCA 467
 Db 323 CCAGGCAATACGTGCTGATCCGATCCGTTATTTAATCCTCTAAACCTGGGTTTGCA 264
 OY 468 ACTGATTAACGCGAAGCTGATGACGCGATCTCTAGACGGGCAAGGTCATTTGCTGA 527
 Db 263 ACTGATTAACGCGAAGCTGATGACGCGATCTCTAGACGGGCAAGGTCATTTGCTGA 204
 OY 528 CTTTACCGGGCATTCGGCAAAACGGCGTTTCGCGAATGAAACGGGTCTTAATTTTCGCA 587
 Db 203 CTTTACCGGGCATTCGGCAAAACGGCGTTTCGCGAATGAAACGGGTCTTAATTTTCGCA 144
 OY 588 ATCAAACTTGTGCTTTAAACGTGAGAAACAGAACGAAAGCTGTTCAATTAACGAGCATT 647
 Db 143 ATCAAACTTGTGCTTTAAACGTGAGAAACAGAACGAAAGCTGTTCAATTAACGAGCATT 84
 OY 648 ACCATCGGAACCTCAAGTGAAGCGCGAACAATGTCATTAAACCGGTGCGTAAAGCTCGC 707
 Db 83 ACCATCGGAACCTCAAGTGAAGCGCGAACAATGTCATTAAACCGGTGCGTAAAGCTCGC 24
 OY 708 ATCAATGCTGACGAGATATTTC 730
 Db 23 ATCAATGCTGACGAGATATTTC 1

RESULT 12
 LOCUS BU713769 531 bp mRNA linear EST 23-OCT-2003
 DEFINITION SJAABU01 Adult sjc 7/94 Schistosoma japonicum cDNA similar to
 sp|P07102|PpA ECOLI PERIPLASMIC APPA PROTEIN PRECURSOR [INCLUDS:
 PHOSPHOANHYDRIDE PHOSPHOHYDROLASE (PH 2.5 ACID PHOSPHATASE) (AP);
 6-PHITASE J, mRNA sequence.
 ACCESSION BU713769
 VERSION BU713769.1 GI:28321125
 KEYWORDS EST.

SOURCE
 ORGANISM Schistosoma japonicum
 Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
 Strigoida; Schistosomatidae; Schistosomatidae; Schistosoma.
 REFERENCE 1 (bases 1 to 531)
 AUTHORS Hu,W., Yan,Q., Shen,D.K., Liu,F., Zhu,Z.D., Song,H.D., Xu,X.R.,
 Wang,Z.J., Rong,Y.P., Zeng,L.C., Wu,J., Zhang,X., Wang,J.J.,
 Xu,X.N., Wang,S.Y., Fu,G., Zhang,X.L., Wang,Z.Q., Brindley,P.J.,
 McManus,B.P., Xue,C.L., Feng,Z., Chen,Z. and Han,Z.G.
 TITLE Evolutionary and biomedical implications of a Schistosoma japonicum
 JOURNAL complementary DNA resource
 PUBMED Nat. Genet. 35 (2), 139-147 (2003)
 COMMENT Contact: Zeguang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 201203, P. R. China
 Tel: 86-21-50801919 (ex.45)
 Fax: 86-21-50801922
 Email: hanzg@chgc.sh.cn.

FEATURES
 source
 1..531
 /organism="Schistosoma japonicum"
 /mol_type="mRNA"
 /strain="Chinese (Anhui) strain"
 /db_xref="taxon:6182"
 /sex="Male and female"
 /tissue_type="whole body"
 /dev_stage="Adult worms"
 /lab_host="Mouse and rabbit"
 /clone_id="Adult sjc 7/94"
 /note="Vector: Lambda ZAP-III XR.; Site 1: EcoR I; Site 2:
 XhoI I; Several hundred adult Schistosoma japonicum
 (Anhui, P.R. China, strain), of mixed sex, were perfused
 from the mesenteries of experimentally infected mice and
 rabbits at the Queensland Institute of Medical Research,
 Brisbane, Australia (QIMR), and stored for several months
 in liquid nitrogen. Subsequently, mRNA was isolated at the
 QIMR from lysates of these worms by oligo dT
 chromatography, using a kit from Pharmacia. The mRNA was
 then shipped to Clontech, Palo Alto, CA, USA, who
 constructed a cDNA library. First strand synthesis was
 primed with an oligo-dT-XhoI-primer and synthesized using
 M-MLV reverse transcriptase. Second strand synthesis was
 accomplished with RNase H and T4 DNA polymerase. The
 double stranded cDNA was ligated to EcoRI linkers,
 digested with EcoRI and XhoI, and ligated into the
 phagemid vector lambda ZAP II XR. After construction of
 this directional library by Clontech, it was returned to
 the QIMR. During analysis of the library at the QIMR, we
 have found that a small percentage, 2% to 3%, of the
 clones contain inserts that appear to be highly homologous
 to sequences from salmonid fishes, as determined by
 homology comparisons using BLAST and by Southern
 hybridization analysis to genomic DNA from salmon (Sigma
 Chemical Co., St. Louis, MO) under stringent washing
 conditions. The remainder of the clones appear to contain
 S. japonicum sequences."

ORIGIN
 Query Match 26.4%; Score 348.6; DB 3; Length 531;
 Best Local Similarity 88.0%; Pred. No. 7.9e-97;
 Matches 425; Conservative 0; Mismatches 54; Indels 4; Gaps 4;

OY 670 GCCGCAATGCTCATTAACCGGTGCGGTAAAGCTGCGATCAATGCTGACGAGATATT 729
 Db 1 GCCGCAATGCTCATTAACCGGTGCGGTAAAGCTGCGATCAATGCTGACGAGATGGG 60
 OY 730 CTCTCGCAACAGGACGAGGGAATCCCGGAGCGGGGTGGGGAAGATCAAGATTCAAC 789
 Db 61 CTCTCGCAACAGGACGAGGGAATCCCGGAGCGGGGTGGGGAAGATCAAGATTCAAC 120
 OY 790 CAGTGAACACCTTGCTAAGTTGCAATTAACCGCAATTTATTGCTAACAGCAGCGCA 849

```

Db      121  CAGTGAAACCTTGTGTAAGTTGCAATACGCGCAATTTTATTTGCTACAGACGCGCA 180
Qy      850  GAGGTGCCCCGAGCCGCGCCCAACCCCGTTATTTGATTTGATCATGCGAGCGTTAGCCCC 909
Db      181  GAGGTGCCCCGAGCCGCGCCCAACCCCGTTATTTGATTTGATCATGCGAGCGTTAGCCCC 240
Qy      910  CATCCACCGCAAAAACAGCGGTATGTTGATGACATTAACCACTTCACT-ACGTGTTATTCG 968
Db      241  CGTTATCATTTATGATTTGATGATGAGAGACATTCCTCTTGAATGACGAAAAAAG 300
Qy      969  CGGACACGATCTAATCTGCAAAATCTCGCGCGGACCTGAGACTCACTGACGCTTC 1028
Db      301  CGGCGCAGAGCTGAGCTGATTAACCTCGCGCGGACGAGCTCACTGAGCGCTGTC 360
Qy      1029  CGGTACCGCGGATTAACCGCGCGGAGGTGTAACCTG-TCGTTGAACGCT-GGCGTCG 1086
Db      361  CAGGCGCGCGGATTAACCGCGCGGAGCGGAACTGGAATGTTGAACGCTGCGCGCTCG 420
Qy      1087  CTAAAGGATTAACGAGCGGATGATTCAGGTTGCTGCTTCGACTTACAGCAGATG 1146
Db      421  CTAAAGGATTAACGAGCGGATGATTCAGGTTGCTGCTTCGAGC-TTACAGCAGATG 479
Qy      1147  CGT 1149
Db      480  CGT 482

```

```

Matches 335; Conservative 0; Mismatches 16; Indels 2; Gaps 2;
Qy      497  TCTCAGACGAGGCGAGAGGCTAATTGCTGACTTTACCGGCGCATCGGCAACGGCGTTTC 556
Db      351  TCTTACGACGAGGCGAGAG-GGTCAATTGCTGATTTACCGGCGCATCGGCAACGGCGTTTC 293
Qy      557  GCGAAGTGAACGAGGCGCTAATTTTCCGCAATCAAACTGAGCTTTAAACGTAAGAAAC 616
Db      292  GCGAAGTGAACGAGGCTG-TTAAATTTTCCGCAATCAAACTGAGCTTTAAACGTAAGAAAC 234
Qy      617  AGAGAGAAAGCTGTTCAATTAAACGAGGCAATTAACATCGGAATCAAGTGAAGCGCGAGA 676
Db      233  AGAGAGAAAGCTGTTCAATTAAACGAGGCAATTAACATCGGAATCAAGTGAAGCGCGAGA 174
Qy      677  ATGTCATTAACCGGTGCGGTAAAGCTTCGCAATCAATGCTGAACGAGATATTTCTCTGC 736
Db      173  ATGTCATTAACCGGTGCGGTAAAGCTTCGCAATCAATGCTGAACGAGATATTTCTCTGC 114
Qy      737  AACAGGACAGGGAATGCGCGAGCGGCGGAGGAGAAAGATCAACGATTCACACAGTGA 796
Db      113  AACAGGACAGGGAATGCGCGAGCGGCGGAGGAGAAAGATCAACGATTCACACAGTGA 54
Qy      797  ACACCTTGCTAAGTTGCTAATTAACGAGGCAATTTATTTGCTACAGCAGCGCA 849
Db      53  ACACCTTGCTAAGTTGCTAATTAACGAGGCAATTTATTTGCTACAGCAGCGCA 1

```

```

RESULT 14
AA545747/c 706 bp mRNA linear EST 20-MAY-2004
LOCUS CN759004
DEFINITION ID0AA24BC04B01 ApMs Acyrthosiphon pisum cDNA clone ID0AA24BC04
5', mRNA sequence.
ACCESSION CN759004
VERSION CN759004.1 GI:47532927
KEYWORDS EST.
SOURCE Acyrthosiphon pisum (pea aphid)
ORGANISM Acyrthosiphon pisum
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pserygota;
Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.
REFERENCE 1 (bases 1 to 706)
AUTHORS Hunter, W., Martinez-Torres, D., Rahbe, Y., Sabater-Munoz, B.,
Stern, D., Tagu, D. and Wincker, P.
An expressed sequence tags database for the pea aphid Acyrthosiphon
pisum
JOURNAL Unpublished (2004)
COMMENT Contact: D. Tagu
INRA Rennes
UMR BIO3P, BP 35327, F-35653 Le Rheu Cedex France
Tel: +33.2.23.48.51.65
Fax: +33.2.23.48.51.50
Risk of contamination by bacterial sequences from obligatory
(Buchner) or facultative endosymbionts. These sequences were
obtained in the frame of the International Consortium of Aphid
Genomics in collaboration with Genoscope
PCR PRIMERS
FORWARD: CAGGAAACAGCTATGACC
Plate: 24 row: C column: 4.
FEATURES
source 1..706
location/Qualifiers
1..354
/organism="Acyrtosiphon pisum"
/mol_type="mRNA"
/db_xref="taxon:7029"
/cultivar="developmentstage"
/db_xref="taxon:7029"
/clone="ID0AA24BC04"
/tissue_type="whole insect"
/dev_stage="nymphs and adults (parthenogenetic females)"
/lab_host="XLI-Blue"
/clone_lib="ApMs"
/notes="Vector: pBS-SK minus; Site_1: EcoRI; Site_2: XhoI;
Sample name: ID0AA ; plant growth place: Department of
Ecology & Evolutionary Biology, Princeton University ;
Soil conditions: Soil ; Sowing date: 01/06/1999 ;

```

```

RESULT 13
AA545747/c 354 bp mRNA linear EST 12-MAY-1999
LOCUS HBMSF184-REV Human Bone Marrow Stromal Fibroblast Homo sapiens cDNA
DEFINITION clone HBMSF184 5', mRNA sequence.
ACCESSION AA545747
VERSION AA545747.1 GI:2307026
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 354)
AUTHORS Jia, L., Robey, P., Young, M., Schuler, G., Powell, J., Yang, L.,
Lennon, G., Hillier, L., Allen, M., Bowles, L., Geisel, S., Kucaba, T.,
Marras, M., Martin, J., Steptoe, M., Tan, P., Theising, B., Bowers, Y.,
Wylie, F., Waterston, R., Wilson, R. and Francomano, C.
WashU-MGB/NHGRI EST Project
Unpublished (1997)
COMMENT Contact: Libin Jia
Medical Genetics Branch
National Human Genome Research Institute
10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
Tel: 301-402-4877
Fax: 301-496-7157
Email: libin@helix.nih.gov
Seq primer: M13 Reverse
FEATURES
source 1..354
location/Qualifiers
1..354
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HBMSF184"
/sex="Male and Female"
/tissue_type="bone marrow"
/cell_type="stromal fibroblast"
/dev_stage="mixed"
/lab_host="XLI-Blue"
/clone_lib="Human Bone Marrow Stromal Fibroblast"
/notes="Vector: pBluescript; Site_1: EcoRI; Site_2: XhoI"
ORIGIN
Query Match 23.0%; Score 304.6; DB 1; Length 354;
Best Local Similarity 94.9%; Pred. No. 3.9e-83;

```

```

RESULT 14
CN759004/c 706 bp mRNA linear EST 20-MAY-2004
LOCUS CN759004
DEFINITION ID0AA24BC04B01 ApMs Acyrthosiphon pisum cDNA clone ID0AA24BC04
5', mRNA sequence.
ACCESSION CN759004
VERSION CN759004.1 GI:47532927
KEYWORDS EST.
SOURCE Acyrthosiphon pisum (pea aphid)
ORGANISM Acyrthosiphon pisum
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pserygota;
Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.
REFERENCE 1 (bases 1 to 706)
AUTHORS Hunter, W., Martinez-Torres, D., Rahbe, Y., Sabater-Munoz, B.,
Stern, D., Tagu, D. and Wincker, P.
An expressed sequence tags database for the pea aphid Acyrthosiphon
pisum
JOURNAL Unpublished (2004)
COMMENT Contact: D. Tagu
INRA Rennes
UMR BIO3P, BP 35327, F-35653 Le Rheu Cedex France
Tel: +33.2.23.48.51.65
Fax: +33.2.23.48.51.50
Risk of contamination by bacterial sequences from obligatory
(Buchner) or facultative endosymbionts. These sequences were
obtained in the frame of the International Consortium of Aphid
Genomics in collaboration with Genoscope
PCR PRIMERS
FORWARD: CAGGAAACAGCTATGACC
Plate: 24 row: C column: 4.
FEATURES
source 1..706
location/Qualifiers
1..354
/organism="Acyrtosiphon pisum"
/mol_type="mRNA"
/db_xref="taxon:7029"
/cultivar="developmentstage"
/db_xref="taxon:7029"
/clone="ID0AA24BC04"
/tissue_type="whole insect"
/dev_stage="nymphs and adults (parthenogenetic females)"
/lab_host="XLI-Blue"
/clone_lib="ApMs"
/notes="Vector: pBS-SK minus; Site_1: EcoRI; Site_2: XhoI;
Sample name: ID0AA ; plant growth place: Department of
Ecology & Evolutionary Biology, Princeton University ;
Soil conditions: Soil ; Sowing date: 01/06/1999 ;

```

Harvesting date: 01/06/1999 ; Stress date: no stress ;
Description: Aphids inoculated on one-week old *Vicia faba*
under non-sterile conditions. All parthenogenetic stages
and both winged and wingless adults were collected for
library construction ; experimental condition: long
photoperiod (16-hr light/8-hr dark at 18 °C)

ORIGIN

Query Match 22.4% ; Score 295.8 ; DB 8 ; Length 706 ;
Best Local Similarity 99.3% ; Pred. No. 2.6e-80 ;
Matches 297 ; Conservative 0 ; Mismatches 2 ; Indels 0 ; Gaps 0 ;

QY 1002 CGCACTGGAGCTCACTGAGCGCTTCCCGTCAGCCGAGTAACAAGCCCGCAGGTGTGA 1061
DB 706 CGCACTGAGCTCACTGAGCGCTTCCCGTCAGCCGAGTAACAAGCCCGCAGGTGTGA 647
QY 1062 ACTGCTGTTTGAACGCTGCGCTGAGCTAAGCGATTAACGCGATTCAGGTTTGGCT 1121
DB 646 ACTGCTGTTTGAACGCTGCGCTGAGCTAAGCGATTAACGCGATTCAGGTTTGGCT 587
QY 1122 GCTCTTCCAGCTTACAGCAGATGCGATTAACGCGCTGCTATTAATTAACGCGCC 1181
DB 586 GCTCTTCCAGCTTACAGCAGATGCGATTAACGCGCTGCTATTAATTAACGCGCC 527
QY 1182 CGGAGAGGTGAACCTGAGCTGCGAGATGTTGAAGAGCGAATGCGCAGGCGATGTTC 1241
DB 526 CGGAGAGGTGAACCTGAGCTGCGAGATGTTGAAGAGCGAATGCGCAGGCGATGTTC 467
QY 1242 GTTGGCAGGTTTACGCAATGCTGAATGAAGCAGCATACCGGCTGCAAGTTTGAGAT 1300
DB 466 GTTGGCAGGTTTACGCAATGCTGAATGAAGCAGCATACCGGCTGCAAGTTTGATAT 408

RESULT 15

BU619443 707 bp mRNA linear EST 01-OCT-2003
LOCUS BU619443 NIBB Mochii normalized Xenopus early gastrula library
DEFINITION Xenopus laevis cDNA clone XL189d24 5', mRNA sequence.
ACCESSION BU619443
VERSION BU619443.1 GI:37258203
KEYWORDS EST.

ORGANISM Xenopus laevis (African clawed frog)
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Xenopus.

REFERENCE 1 (bases 1 to 707)
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-I,T. and
Kohara,Y.
TITLE Expressed genes in *X. laevis* embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp
The information of this clone is available through the following
URL.
http://xenopus.nibb.ac.jp.

FEATURES
SOURCE location/Qualifiers
1..707

/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="XL189d24"
/tissue_type="whole embryo"
/dev_stage="stage 10.5"
/clone_lib="NIBB Mochii normalized Xenopus early gastrula
library"

ORIGIN

Query Match 18.8% ; Score 248 ; DB 2 ; Length 707 ;
Best Local Similarity 98.0% ; Pred. No. 2.2e-65 ;
Matches 251 ; Conservative 0 ; Mismatches 5 ; Indels 0 ; Gaps 0 ;

QY 750 AATCCCGAGCGCGGGGTTGGGGAAGATCAGGATTTCAACCAAGTGAACACCTTGCTTAG 809
DB 1 AATCCCGAGCGCGGGGTTGGGGAAGATCAGGATTTCAACCAAGTGAACACCTTGCTTAG 60
QY 810 TTTCATTAACCGCGCAATTTTATTTGCTTACAACGACCGCAGAGGTTGCCGACCGCGC 869
DB 61 TTTCATTAACCGCGCAATTTTATTTGCTTACAACGACCGCAGAGGTTGCCGACCGCGC 120
QY 870 CACCCGTTATTTGATTTGATCATGACGCGTTGACCGCCCATCACCGCAAAAACAGGC 929
DB 121 CACCCGTTATTTGATTTGATCATGACGCGTTGACCGCCCATCACCGCAAAAACAGGC 180
QY 930 GTATGTTGACATTACCCACTTCACTGTTTATTTGCGGACAGATTAATCTTGGC 989
DB 181 GTATGTTGACATTACCCACTTCACTGTTTATTTGCGGACAGATTAATCTTGGC 240
QY 990 AATCTCGCGCGCGCA 1005
DB 241 AATCTCGCGCGCGCA 256

Search completed: June 14, 2006, 14:21:54
Job time : 6948.38 secs


```
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2180
; LENGTH: 2879
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-2180

Query Match      2.7%; Score 36.2; DB 7; Length 2879;
Best Local Similarity 51.2%; Pred. No. 0.035; Indels 0; Gaps 0;
Matches 83; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

134 CCAGGCGCCAGCACTGATGAGATGTCACCCAGACGATGCGCAACTGGCCGGTAA 193
pb CCGGGCCCTTGCTATTGAGAAAGAGGTGTCCCAAGGCGCTCCGGCCGCGCAAGT 329
by 194 AACTGGTTGGCTGACACCGCGNGGTGGTAGCTATGCGCTATCTCGGACATTACCAAC 253
Db 328 GCATGTGTGGGGGAGTCCGCCAGCGGGCGCTTATCATCTCGGCGGTGAGTCCGCTC 269
Qy 254 GCCAGCTCTGTGATGACCGACGATGCTGGCGAAAGAGGCT 295
Db 268 GGGAGCGGCGCTTGAGTGCACCGAGCTGAGACCAAGTTGCT 227

RESULT 3
US-10-953-349-34412/c
; Sequence 34412, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34412
; LENGTH: 979
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
US-10-953-349-34412

Query Match      2.6%; Score 35; DB 6; Length 979;
Best Local Similarity 48.5%; Pred. No. 0.048; Indels 0; Gaps 0;
Matches 95; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

163 ACCCCAGACGATGCGCAACTGCGCGTAAACTGGTTGCTGACACCGCGNGTGGT 222
Db 709 AACACCGGGAGAGACCGGGAGACACGCGGGGACCTGGCTGACTGTGAGTGGT 650
Qy 223 GAGCTAATCGCTTATCTCGGACATTACCAAGCGCGAGCTGTGTGAGCCGACGATTTGCTG 282
Db 649 GAGCGGATTCGACGAGACGGGCGCTTCCGCGCGCCGAGGCGGGGGCGGACGCGGCGT 590
Qy 283 GGGAAAAAGGGCTGCCGCGAGTCTGGTCAAGTCCGATATTGCTGATGTGACGAGAGGT 342
Db 589 GCGGCGAGGCGAGTCCGGGACCGCGGTGACGACCGGTAGAGCTGAGACGCGCAAGCCA 530
Qy 343 ACCCGTAAACAAGGCG 358
Db 529 GTCGGTGAGAGAGGCG 514
```

```
RESULT 4
US-10-953-349-11004/c
; Sequence 11004, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 11004
; LENGTH: 2271
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-11004

Query Match      2.5%; Score 33.4; DB 6; Length 2271;
Best Local Similarity 60.4%; Pred. No. 0.3; Indels 0; Gaps 0;
Matches 55; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

1212 TGAAGAGCGAAATGCGAGGCGATGTGTGCTGGCAGGTTTACGCAATCGTAATGA 1271
Db 1734 TGAACCGGAGACATGTTCCGGGTGATATGTAACCTGAGAAAGGCTTCTGCAGATCTTGAAAGA 1675
Qy 1272 AGCAGCATACCGGCGGTGCACTTGAATCT 1302
Db 1674 CCGGAGAAATGATGTTGCTCTTGATCTTT 1644
```

```
RESULT 5
US-11-217-529-79920/c
; Sequence 79920, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 79920
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-79920

Query Match      2.5%; Score 33.2; DB 7; Length 1257;
Best Local Similarity 64.1%; Pred. No. 0.24; Indels 28; Gaps 0;
Matches 50; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

671 CCGACATGTTTCATTAAACGGTGGTAAAGCTCGCATCATGCTGACGAGATATTTTC 730
Db 253 CCGAGTGTGTCTCAAAAATGGGTTTGTGAGCGTGCATGAGCGCTGAGAGAGCGGTATT 194
Qy 731 TCTGCAACAAGCAAGG 748
Db 193 TCATCCAATCCGTAGAGG 176
```

```
RESULT 6
US-11-293-654-37
```

```
/ Sequence 37, Application US/11293654
/ Publication No. US20060123568A1
/ GENERAL INFORMATION:
/ APPLICANT: The Government of the United States of America, as
/ APPLICANT: Represented by the Secretary of the Dept. of Health and Human
/ APPLICANT: Services
/ APPLICANT: Patton, John T.
/ APPLICANT: Tardorewala, Zenobia F.
/ TITLE OF INVENTION: ANTIGEN DELIVERY PLATFORM
/ FILE REFERENCE: 4239-69406-02
/ CURRENT APPLICATION NUMBER: US/11/293,654
/ CURRENT FILING DATE: 2005-12-02
/ PRIOR APPLICATION NUMBER: US 60/633,036
/ PRIOR FILING DATE: 2004-12-03
/ NUMBER OF SEQ ID NOS: 37
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 37
/ LENGTH: 3431
/ TYPE: DNA
/ ORGANISM: Artificial
/ FEATURE:
/ OTHER INFORMATION: Vector
US-11-293-654-37
```

```
Query Match          2.4%; Score 32.2; DB 7; Length 3431;
Best Local Similarity 66.7%; Pred. No. 1;
Matches 46; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
```

```
QY      1255 ACCGAATGCTGATGATGACGACGATACCGGCTGAGTTGATCTCATCCATCCAC 1314
DB      85  ACAGATTCTATTAAGAGAGAAATTAACCATGGAGGATCCAGATCTCATCCATCCAC 144
QY      1315 CATCACTAA 1323
DB      145 CATCACTAA 153
```

```
RESULT 7
US-10-449-902-8987
/ Sequence 8987, Application US/10449902
/ Publication No. US20060123505A1
/ GENERAL INFORMATION:
/ APPLICANT: National Institute of Agricultural Sciences.
/ APPLICANT: Bio-oriented Technology Research Advancement Institution.
/ APPLICANT: The Institute of Physical and Chemical Research.
/ APPLICANT: Foundation for Advancement of International Science.
/ TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USBS THEREOF
/ FILE REFERENCE: MOA-A0205Y1-US
/ CURRENT APPLICATION NUMBER: US/10/449,902
/ CURRENT FILING DATE: 2003-05-29
/ PRIOR APPLICATION NUMBER: JP 2002-203269
/ PRIOR FILING DATE: 2002-05-30
/ PRIOR APPLICATION NUMBER: JP 2002-383870
/ PRIOR FILING DATE: 2002-12-11
/ NUMBER OF SEQ ID NOS: 56791
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 8987
/ LENGTH: 1396
/ TYPE: DNA
/ ORGANISM: Oryza sativa
/ PUBLICATION INFORMATION:
/ DATABASE ACCESSION NUMBER: AK106335
/ DATABASE ENTRY DATE: 2002-08-28
US-10-449-902-8987
```

```
Query Match          2.4%; Score 32; DB 6; Length 1396;
Best Local Similarity 58.3%; Pred. No. 0.68;
Matches 56; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
```

```
QY      679 GTCCTATTACCGGCGGTGAAGCTCCGATCAATGCTACGAGATATTCTCTCGCAA 738
DB      456 GTCTGTTCTCGGTGCGGCTTCCTCTCGCGAAGCGGAGAGAGCTCTACTCCGAGAA 515
```

```
QY      739 CAAGCAGAGGAATGCCGAGCCGGGTGGGAGAG 774
DB      516 CCGGATGAGGAGCAGGCGGCGCAGCATGTGCGGG 551
```

```
RESULT 8
US-10-449-902-6188/c
/ Sequence 6188, Application US/10449902
/ Publication No. US20060123505A1
/ GENERAL INFORMATION:
/ APPLICANT: National Institute of Agricultural Sciences.
/ APPLICANT: Bio-oriented Technology Research Advancement Institution.
/ APPLICANT: The Institute of Physical and Chemical Research.
/ APPLICANT: Foundation for Advancement of International Science.
/ TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USBS THEREOF
/ FILE REFERENCE: MOA-A0205Y1-US
/ CURRENT APPLICATION NUMBER: US/10/449,902
/ CURRENT FILING DATE: 2003-05-29
/ PRIOR APPLICATION NUMBER: JP 2002-203269
/ PRIOR FILING DATE: 2002-05-30
/ PRIOR APPLICATION NUMBER: JP 2002-383870
/ PRIOR FILING DATE: 2002-12-11
/ NUMBER OF SEQ ID NOS: 56791
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 6188
/ LENGTH: 1489
/ TYPE: DNA
/ ORGANISM: Oryza sativa
/ PUBLICATION INFORMATION:
/ DATABASE ACCESSION NUMBER: AK104957
/ DATABASE ENTRY DATE: 2002-08-28
US-10-449-902-6188
```

```
Query Match          2.4%; Score 31.6; DB 6; Length 1489;
Best Local Similarity 51.9%; Pred. No. 0.98;
Matches 70; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
```

```
QY      213 GCGNGTGTGATGACATATGCTATCTCGAATTAACCAAGCCGACGCTGTGATACCA 272
DB      538 GCGGGGATGTCAGCGCACCGCCATGGCGAAGCTCACCCGCGCAGCTGACGCC 479
QY      273 CGGATTGCTGCGGAAAGAGGCTGCCCGGAGTGTGATGCTGCGATATTGCTGATGT 332
DB      478 CCGCAGACACCTTCTCCAGCTGGGCGCTCACCAACATCAAGTGTCTCTCGTGTGT 419
QY      333 CGACGAGCGTACCG 347
DB      418 CGTCGTCTTCTCCG 404
```

```
RESULT 9
US-10-449-902-17204/c
/ Sequence 17204, Application US/10449902
/ Publication No. US20060123505A1
/ GENERAL INFORMATION:
/ APPLICANT: National Institute of Agricultural Sciences.
/ APPLICANT: Bio-oriented Technology Research Advancement Institution.
/ APPLICANT: The Institute of Physical and Chemical Research.
/ APPLICANT: Foundation for Advancement of International Science.
/ TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USBS THEREOF
/ FILE REFERENCE: MOA-A0205Y1-US
/ CURRENT APPLICATION NUMBER: US/10/449,902
/ CURRENT FILING DATE: 2003-05-29
/ PRIOR APPLICATION NUMBER: JP 2002-203269
/ PRIOR FILING DATE: 2002-05-30
/ PRIOR APPLICATION NUMBER: JP 2002-383870
/ PRIOR FILING DATE: 2002-12-11
/ NUMBER OF SEQ ID NOS: 56791
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 17204
/ LENGTH: 1600
/ TYPE: DNA
/ ORGANISM: Oryza sativa
```

;; PUBLICATION INFORMATION:
;; DATABASE ACCESSION NUMBER: AK067664
;; DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-17204

Query Match 2.4%; Score 31.6; DB 6; Length 1600;
Best Local Similarity 51.9%; Pred. No. 1;
Matches 70; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 213 GCGGNGTGTGAGTAACTGCTATCTGAGACATTACCAAGCCAGCGCTGTGTAGCCGA 272
DB 615 GCGGCACTGCAAGCCAGCCGCAATGCGAAGCTCACCCGCGCCAGCCAGCTGAGCGC 556
QY 273 CGGATTTGCGGAAAAAGGGCTGCGCGGAGTCTGTCAAGTGGCGATTATTGTGATGT 332
DB 555 GCGCGAGCACTTCTCCAGCTGCGGCTCACCAACATAGTGTCTCTCGTGTGTGT 496
QY 333 CGACGAGCGTACCCG 347
DB 495 CGTGCTCTTCTCCG 481

RESULT 10
US-10-449-902-13863/c
; Sequence 13863, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:

;; APPLICANT: National Institute of Agrobiological Sciences.
;; APPLICANT: Bio-oriented Technology Research Advancement Institution.
;; APPLICANT: The Institute of Physical and Chemical Research.
;; APPLICANT: Foundation for Advancement of International Science.
;; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
;; FILE REFERENCE: MOA-A0205Y1-US
;; CURRENT APPLICATION NUMBER: US/10/449,902
;; CURRENT FILING DATE: 2003-05-29
;; PRIOR APPLICATION NUMBER: JP 2002-203269
;; PRIOR FILING DATE: 2002-05-30
;; PRIOR APPLICATION NUMBER: JP 2002-383870
;; PRIOR FILING DATE: 2002-12-11
;; NUMBER OF SEQ ID NOS: 56791
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 13863
;; LENGTH: 3129
;; TYPE: DNA
;; ORGANISM: Oryza sativa
;; PUBLICATION INFORMATION:
;; DATABASE ACCESSION NUMBER: AK111211
;; DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-13863

Query Match 2.4%; Score 31.4; DB 6; Length 3129;
Best Local Similarity 47.7%; Pred. No. 1.8; Indels 0; Gaps 0;
Matches 92; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 999 GCGGCACTGAGCTCAACTGAGCGCTTCCCGTACGCGGATTAACAGCCGCGCAGGTGG 1058
DB 2942 CGAAGTCCGAGGCTCAAGCTGCGCTGCGCAGTTCGTTGCCCGCGTGAATGAGG 2883
QY 1059 TGAAGTGTGTTGAACGCTGCGCTGAGCATTAACAGCCAGGTGATTCAGTTTC 1118
DB 2882 TGCCTGGAACGGTTACCGAGTTTGAAGCCAGAGGAGTGAATGCGCTGCGAACCCATC 2823
QY 1119 GCTGCTTTCAGACTTACAGCAGATGCGTGAATAACGCGCGTGTCAATTAAATACGCC 1178
DB 2822 GGTGAACCCGAGAGATTTGGAGAGCATGAGAGGAAAGCGCCGCGCCATTGCGAGGGA 2763
QY 1179 GCCCGAGAGGTG 1191
DB 2762 ACCAGAGGTGCTG 2750

RESULT 11
US-10-449-902-12516/c

; Sequence 12516, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:

;; APPLICANT: National Institute of Agrobiological Sciences.
;; APPLICANT: Bio-oriented Technology Research Advancement Institution.
;; APPLICANT: The Institute of Physical and Chemical Research.
;; APPLICANT: Foundation for Advancement of International Science.
;; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
;; FILE REFERENCE: MOA-A0205Y1-US
;; CURRENT APPLICATION NUMBER: US/10/449,902
;; CURRENT FILING DATE: 2003-05-29
;; PRIOR APPLICATION NUMBER: JP 2002-203269
;; PRIOR FILING DATE: 2002-05-30
;; PRIOR APPLICATION NUMBER: JP 2002-383870
;; PRIOR FILING DATE: 2002-12-11
;; NUMBER OF SEQ ID NOS: 56791
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 12516
;; LENGTH: 2144
;; TYPE: DNA
;; ORGANISM: Oryza sativa
;; PUBLICATION INFORMATION:
;; DATABASE ACCESSION NUMBER: AK109864
;; DATABASE ENTRY DATE: 2002-08-28
US-10-449-902-12516

Query Match 2.4%; Score 31.2; DB 6; Length 2144;
Best Local Similarity 48.6%; Pred. No. 1.7; Indels 89; Gaps 0;
Matches 84; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 211 CCGGNGTGTGAGTAACTGCTATCTCGACATTACCAAGCCAGCGTGTGTAGCC 270
DB 1652 GCGGGGGGTGGGCAACAGCTGCTCAGCGGAGGATCCATCCCTTCTCTGTCGCG 1593
QY 271 GACGATTTCTGGGAAAAAGGCTGCCCGCAGTCTGTCAAGTTCGGATTAATGCTGAT 330
DB 1592 GTTGTAGGTGTTCCCGCGATGCGCACCTTCTTCTCTCAAGCTTCCGTTGGCGAGAA 1533
QY 331 GTCGACGAGCGTACCCGTAATAACAGGCAAGCTTTCGCGCGGAGGTGCGAC 383
DB 1532 GTCAACCAAGTCTCTGTTGTTCCGAGCGCGCGCGGCGGCGTCCAGC 1480

RESULT 12
US-10-449-902-10760/c
; Sequence 10760, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:

;; APPLICANT: National Institute of Agrobiological Sciences.
;; APPLICANT: Bio-oriented Technology Research Advancement Institution.
;; APPLICANT: The Institute of Physical and Chemical Research.
;; APPLICANT: Foundation for Advancement of International Science.
;; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
;; FILE REFERENCE: MOA-A0205Y1-US
;; CURRENT APPLICATION NUMBER: US/10/449,902
;; CURRENT FILING DATE: 2003-05-29
;; PRIOR APPLICATION NUMBER: JP 2002-203269
;; PRIOR FILING DATE: 2002-05-30
;; PRIOR APPLICATION NUMBER: JP 2002-383870
;; PRIOR FILING DATE: 2002-12-11
;; NUMBER OF SEQ ID NOS: 56791
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 10760
;; LENGTH: 1344
;; TYPE: DNA
;; ORGANISM: Oryza sativa
;; PUBLICATION INFORMATION:
;; DATABASE ACCESSION NUMBER: AK108108
;; DATABASE ENTRY DATE: 2002-08-28
US-10-449-902-10760

Query Match 2.3%; Score 31; DB 6; Length 1344;
Best Local Similarity 53.8%; Pred. No. 1.5;

Matches 64; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 258 GCCTGTGTGTCGACGAGATTGCTGGGAAAAAGGCTGCCGACGCTGTGTCAGTGC 317
DB 289 GCCTGTGTGTCGAGTGAAGCCGTCAGCGCGCGCTCGGCGCGCGCTGTGTGTCAGTGA 230

QY 318 GATTATTTGCTGATGTCGACGAGCGTACCCGTAATAAGGAGCCCTTGCCGCGCGG 376
DB 229 TGTGTGTCGCTGTGTGTCGACGAGGAGTGCCTGACGAAAGAGCCGCTCGCGGAC 171

RESULT 13

US-11-293-697-707/c
; Sequence 707, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; PRIOR FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 707
; LENGTH: 2791
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-707

Query Match

Best Local Similarity 2.3%; Score 31; DB 7; Length 2791;
Matches 61; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 59 CATTGCTCAGATGAGCCGAGCTGAGCTGGAAGTGGTATTCAGTCGTGATG 118
DB 1419 CCTTGCACTGAGTGTGTACAGCTCTTAAGCGCGTGTCTGAGTTGTTCTCTCTG 1360

QY 119 GTGTGCTGCTCCACCAAGCCGCAACTGATGCAAGATGTCACCCGAG 169
DB 1359 GTGGGTTCGTGTGCTGCTGCTGCTCAGAGTGAAGTGAAGTCTTCGCGG 1309

RESULT 14

US-10-953-349-27515/c
; Sequence 27515, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 27515
; LENGTH: 1393
; TYPE: DNA
; ORGANISM: Trifolium aestivum
US-10-953-349-27515

Query Match

Best Local Similarity 2.3%; Score 30.8; DB 6; Length 1393;
Matches 89; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 603 TAAAGTGAGAAACGAGAAAGCTGTTCATTACGCAAGGATTCACCTGCAACTCAA 662
DB 1167 TAACTTTTGTGATCTTGTGAGGTTCTTACCAACAGGTAAAGCTTACAGCTACAGTCCA 1108

QY 663 GGTGAGCGCCGCAATGTCTCTATTACCGCGTGAAGCCTGCATCAATGCTGACGGA 722

DB 1107 GGGGTGGGCGCGGGGGGTGATGCAACAGCTTACATTCAGTGTCTCCCGG 1048

QY 723 GATATTTCTCTGCAACAAACACAGGAAATGCCGAGCCGGGTGGGAAAGATCACCGA 782
DB 1047 GATTAGTCACTACATACATCAACAGCGTCCGTGTGTAAGACAGGAAAGAGTCTCACAGA 988

QY 783 TTCACA 788
DB 987 TTTACA 982

RESULT 15

US-10-449-902-11476/c
; Sequence 11476, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agricultural Sciences
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11476
; LENGTH: 1461
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ENTRY DATE: 2002-08-28
US-10-449-902-11476

Query Match

Best Local Similarity 2.3%; Score 30.8; DB 6; Length 1461;
Matches 65; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 253 GCGCAGCTCTGTGACGAGCGATTCGTGCGAAAAAGGCTGCCGACGCTGCTCAG 312
DB 1072 CGCGACGCGCGGTTGACAGCAGACACTGTGCTGTCTGCGCGGAGTGGGTGAG 1013

QY 313 GTGCGATTATTCGATGATGAGAGGAGGTCACCCGTAATAACAGGAGGAGCCCTTCGCCGCG 372
DB 1012 CAGCGCAATCTGTGCTCATGTGTGACGAGAGACGATGAAGTCGGGGAAGAACTGTCGTG 953

QY 373 GG 374
DB 952 GG 951

Search completed: June 14, 2006, 14:23:15
Job time : 41.2281 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2006, 10:29:05 ; Search time 1615.16 Seconds
(without alignments)
10064.996 Million cell updates/sec

Title: US-10-601-319-1

Perfect score: 1322

Sequence: 1 atgaagaagatcttaacc.....atcaccatcaccatcactaa 1323

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main:

1: /EMC_Celerra_SIDS3/ptodata/2/pubna/US07_PUBCOMB.seq:*
2: /EMC_Celerra_SIDS3/ptodata/2/pubna/US08_PUBCOMB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/2/pubna/US09_PUBCOMB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/2/pubna/US09B_PUBCOMB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/2/pubna/US09C_PUBCOMB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/2/pubna/US10A_PUBCOMB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/2/pubna/US10B_PUBCOMB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/2/pubna/US10C_PUBCOMB.seq:*
9: /EMC_Celerra_SIDS3/ptodata/2/pubna/US10D_PUBCOMB.seq:*
10: /EMC_Celerra_SIDS3/ptodata/2/pubna/US10E_PUBCOMB.seq:*
11: /EMC_Celerra_SIDS3/ptodata/2/pubna/US10F_PUBCOMB.seq:*
12: /EMC_Celerra_SIDS3/ptodata/2/pubna/US10G_PUBCOMB.seq:*
13: /EMC_Celerra_SIDS3/ptodata/2/pubna/US11A_PUBCOMB.seq:*
14: /EMC_Celerra_SIDS3/ptodata/2/pubna/US11B_PUBCOMB.seq:*
15: /EMC_Celerra_SIDS3/ptodata/2/pubna/US11C_PUBCOMB.seq:*
16: /EMC_Celerra_SIDS3/ptodata/2/pubna/US11D_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1322	100.0	1323	US-09-777-566A-1	Sequence 1, Appl1
2	1322	100.0	1323	US-09-866-379-1	Sequence 1, Appl1
3	1322	100.0	1323	US-10-034-985-1	Sequence 1, Appl1
4	1322	100.0	1323	US-10-430-356-1	Sequence 1, Appl1
5	1322	100.0	1323	US-10-601-319-1	Sequence 1, Appl1
6	1322	100.0	1323	US-10-933-115-1	Sequence 1, Appl1
7	1287.8	97.4	1301	US-09-866-379-7	Sequence 7, Appl1
8	1287.8	97.4	1301	US-10-156-660-3	Sequence 7, Appl1
9	1287.8	97.4	1301	US-10-601-319-7	Sequence 7, Appl1
10	1287.8	97.4	1301	US-10-933-115-7	Sequence 7, Appl1
11	1287.8	97.4	1301	US-11-056-354-3	Sequence 3, Appl1
12	1287.8	97.4	1301	US-10-282-122A-7167	Sequence 7167, Ap
13	1286.2	97.3	1301	US-09-866-379-5	Sequence 5, Appl1
14	1286.2	97.3	1301	US-09-866-379-9	Sequence 9, Appl1
15	1284.6	97.2	1301	US-10-601-319-5	Sequence 5, Appl1
16	1284.6	97.2	1301	US-10-933-115-5	Sequence 5, Appl1
17	1283	97.0	1301	US-09-866-379-6	Sequence 6, Appl1

18	1283	97.0	1301	US-10-601-319-6	Sequence 6, Appl1
19	1283	97.0	1301	US-10-933-115-6	Sequence 6, Appl1
20	1265	95.7	1308	US-10-156-660-1	Sequence 1, Appl1
21	1265	95.7	1308	US-10-601-319-9	Sequence 9, Appl1
22	1265	95.7	1308	US-10-933-115-9	Sequence 9, Appl1
23	1263.8	95.6	1489	US-10-266-041-9	Sequence 9, Appl1
24	1263.8	95.6	1489	US-10-284-962-1	Sequence 1, Appl1
25	1263	95.5	1486	US-11-018-709-2	Sequence 2, Appl1
26	1255	94.9	1396	US-11-056-354-1	Sequence 4, Appl1
27	1251.8	94.7	1486	US-10-284-962-4	Sequence 4, Appl1
28	1251.8	94.7	1486	US-11-018-709-4	Sequence 4, Appl1
29	728.4	55.1	11357	US-11-074-522-14	Sequence 14, Appl1
30	689.4	52.1	13737	US-10-450-763-25689	Sequence 25689, A
31	666.4	50.4	1281	US-10-334-672-4	Sequence 4, Appl1
32	666.4	50.4	1281	US-10-334-671-4	Sequence 4, Appl1
33	399.4	30.2	4050	US-10-450-763-386	Sequence 386, App
34	399	30.2	466	US-10-282-122A-3181	Sequence 3181, Ap
35	390	29.5	1264	US-10-450-763-30120	Sequence 30120, A
36	390	29.5	1264	US-10-450-763-30120	Sequence 30120, A
37	222.8	16.9	1326	US-10-282-122A-41608	Sequence 41608, A
38	218.8	16.6	1326	US-10-021-723A-3	Sequence 3, Appl1
39	215.6	16.3	1325	US-10-021-723A-1	Sequence 1, Appl1
40	206.6	15.6	1326	US-10-021-723A-11	Sequence 11, Appl1
41	199	15.1	356	US-10-450-763-25687	Sequence 25687, A
42	164.2	12.4	254	US-10-317-444-301	Sequence 301, App
43	164.2	12.4	254	US-10-317-444-302	Sequence 302, App
44	164.2	12.4	254	US-10-317-444-303	Sequence 303, App
45	164.2	12.4	254	US-10-317-444-304	Sequence 304, App

ALIGNMENTS

RESULT 1
US-09-777-566A-1
Sequence 1, Application US/09777566A
Patent No. US20010055788A1
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: SHORT, Jay
TITLE OF INVENTION: RECOMBINANT BACTERIAL PHRYASIS AND USES THEREOF
FILE REFERENCE: DIVER370-6
CURRENT APPLICATION NUMBER: US/09/777,566A
CURRENT FILING DATE: 2001-06-11
PRIOR APPLICATION NUMBER: US 09/318,528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: US 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 09/259,214
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: US 08/910,798
PRIOR FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 1323
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1320)
NAME/KEY: misc_feature
LOCATION: (1)..(1323)
OTHER INFORMATION: n is any nucleotide
US-09-777-566A-1

Query Match 100.0%; Score 1322; DB 3; Length 1323;
Beat Local Similarity 100.0%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ATGAAGAAGATCTTAATCCATTTTATCTTCGTGATTCGTTAACCCGCAATTCGCA 60
|||||

Db 1 ATGAAAGGATCTTATCCATTTTATCTCTGATTCGTTACCCCGCAATCTGCA 60
QY 61 TTCCGCTCAGAGTGAAGCCGAGAGCTGGAAGTGTGTGATTTGATGTCATATGTT 120
Db 61 TTCCGCTCAGAGTGAAGCCGAGAGCTGGAAGTGTGTGATTTGATGTCATATGTT 120
QY 121 GTGGGTGTCTCAACCAAGGAGCCGACGATGTGAGATGTCAACCCGAGAGCATGGCA 180
Db 121 GTGGGTGTCTCAACCAAGGAGCCGACGATGTGAGATGTCAACCCGAGAGCATGGCA 180
QY 181 ACTGGCCGCTGTAACCTGGGTGGTGTGACCAACCGGNGTGTGTGATGCTATCTC 240
Db 181 ACTGGCCGCTGTAACCTGGGTGGTGTGACCAACCGGNGTGTGTGATGCTATCTC 240
QY 241 GGAATTACCAACGCGCAGGCTGTGGTACCGACGATTTGCTGGCAAAAAGGCTGCCG 300
Db 241 GGAATTACCAACGCGCAGGCTGTGGTACCGACGATTTGCTGGCAAAAAGGCTGCCG 300
QY 301 CAGTCTGGTCAAGGTGCGGATTTATGCTGATGTGCAAGCGTAAACCGTAAACAGGCGA 360
Db 301 CAGTCTGGTCAAGGTGCGGATTTATGCTGATGTGCAAGCGTAAACCGTAAACAGGCGA 360
QY 361 GCTTTCGCGCGCGGCTGGCACTGATGTGCAATTAACCGTACATACCAGGAGATACG 420
Db 361 GCTTTCGCGCGCGGCTGGCACTGATGTGCAATTAACCGTACATACCAGGAGATACG 420
QY 421 TCCAGTCCGATCCGTTATTTATCTCTGATTAACCTGGGCTTGGCACTGATTAACCG 480
Db 421 TCCAGTCCGATCCGTTATTTATCTCTGATTAACCTGGGCTTGGCACTGATTAACCG 480
QY 481 AACGTAAGTCAAGGATCTCTGAGAGGAGGAGGATGCAATTTGATCTTACCGGAGAT 540
Db 481 AACGTAAGTCAAGGATCTCTGAGAGGAGGAGGATGCAATTTGATCTTACCGGAGAT 540
QY 541 CGGCAACGCGCTTTCGCGAATGGAACGCGGTCTTAAATTTTCGCAATCAAACTTGTGC 600
Db 541 CGGCAACGCGCTTTCGCGAATGGAACGCGGTCTTAAATTTTCGCAATCAAACTTGTGC 600
QY 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGAGGAGATTAACATGAGATC 660
Db 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGAGGAGATTAACATGAGATC 660
QY 661 AAGGAGGAGGCGGAGAAATGTCATTAACCGGTGGGTGAGGCTGAGATCAATGTCAGG 720
Db 661 AAGGAGGAGGCGGAGAAATGTCATTAACCGGTGGGTGAGGCTGAGATCAATGTCAGG 720
QY 721 GAGATATTTCTCTCTCAACCAAGCAGGAGATGCGGAGCCGCGGTGGGGAAGATCAAC 780
Db 721 GAGATATTTCTCTCTCAACCAAGCAGGAGATGCGGAGCCGCGGTGGGGAAGATCAAC 780
QY 781 GATTCAACACAGTGAACACCTTGTCTAAGTTTGCATTAACGCGCAATTTTATTTGCTACA 840
Db 781 GATTCAACACAGTGAACACCTTGTCTAAGTTTGCATTAACGCGCAATTTTATTTGCTACA 840
QY 841 CGCAGCGCAGAGGTTGCGCGGAGCGCGCACCCCGTTTGAATTTGATCAAGGAGG 900
Db 841 CGCAGCGCAGAGGTTGCGCGGAGCGCGCACCCCGTTTGAATTTGATCAAGGAGG 900
QY 901 TTGAGCGCCCATCAACCGCAAAAACAGCGATATGTTGATCAATTAACCTTCAAGTCTG 960
Db 901 TTGAGCGCCCATCAACCGCAAAAACAGCGATATGTTGATCAATTAACCTTCAAGTCTG 960
QY 961 TTTATTTGCGGACACGATCTAATCTGCGAAATCTCGCGCGGCGCACTGAGCTCAACTGG 1020
Db 961 TTTATTTGCGGACACGATCTAATCTGCGAAATCTCGCGCGGCGCACTGAGCTCAACTGG 1020
QY 1021 AAGCTTCCCGGTCAGCGGATTAACAGCGGAGGAGGAGTGAACCTGGTGTGTAAGCGTGG 1080
Db 1021 AAGCTTCCCGGTCAGCGGATTAACAGCGGAGGAGGAGTGAACCTGGTGTGTAAGCGTGG 1080
QY 1081 CGTGGCTAAGGATTAACAGCGAGTGAATTCAGTTCGCTGCTTCCAGACTTTACAG 1140
Db 1081 CGTGGCTAAGGATTAACAGCGAGTGAATTCAGTTCGCTGCTTCCAGACTTTACAG 1140

QY 1141 CAGATGGGTGATTAAGCGCGCTGTCAATTAATACGCGCCCGGAGAGTGAACCTGACC 1200
Db 1141 CAGATGGGTGATTAAGCGCGCTGTCAATTAATACGCGCCCGGAGAGTGAACCTGACC 1200
QY 1201 CTGGCAGAGATGTAAGAGCGAAGATGCGCAGGAGATGTTGTTGGCAAGTTTACGCA 1260
Db 1201 CTGGCAGAGATGTAAGAGCGAAGATGCGCAGGAGATGTTGTTGGCAAGTTTACGCA 1260
QY 1261 ATCGTAATGAAGACGATTCGCGCGTGCAGTTTGAATCTCATCAACCATCACTAC 1320
Db 1261 ATCGTAATGAAGACGATTCGCGCGTGCAGTTTGAATCTCATCAACCATCACTAC 1320
QY 1321 TAA 1323
Db 1321 TAA 1323

RESULT 2
US-09-866-379-1
Sequence 1, Application US/09866379
Patent No. US20020136754A1
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: SHORT, Jay
APPLICANT: KRETZ, Keith
APPLICANT: GRAY, Kevin
APPLICANT: BARTON, Nelson
APPLICANT: GARRETT, James
APPLICANT: O'DONOHUE, Bileen
TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
FILE REFERENCE: DIVER1370-7
CURRENT APPLICATION NUMBER: US/09/866,379
CURRENT FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: US 09/580,515
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 09/318,528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: US 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 09/259,214
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: US 08/910,798
PRIOR FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1323
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)..(1323)
OTHER INFORMATION: n is any nucleotide
NAME/KEY: CDS
LOCATION: (1)..(1323)
OTHER INFORMATION:
US-09-866-379-1

Query Match 100.0%; Score 1323; DB 3; Length 1323;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAGGATCTTATCCATTTTATCTCTGATTCGTTACCCCGCAATCTGCA 60
Db 1 ATGAAAGGATCTTATCCATTTTATCTCTGATTCGTTACCCCGCAATCTGCA 60
QY 61 TTCCGCTCAGAGTGAAGCCGAGAGCTGGAAGTGTGTGATTTGATGTCATATGTT 120
Db 61 TTCCGCTCAGAGTGAAGCCGAGAGCTGGAAGTGTGTGATTTGATGTCATATGTT 120
QY 121 GTGGGTGTCTCAACCAAGGAGCCGACGATGTGAGATGTCAACCCGAGAGCATGGCA 180
Db 121 GTGGGTGTCTCAACCAAGGAGCCGACGATGTGAGATGTCAACCCGAGAGCATGGCA 180

Db 121 GTGCGGTCTCAACCAAGGCGACGCACTGATGACAGATGTCAACCCAGACGATGCGCA 180
 Qy 181 ACTTGCGCGGTAAATCTGGGTGGCTGACACCGCGAGGTGGAGTAAATCGCTATCTC 240
 Db 181 ACTTGCGCGGTAAATCTGGGTGGCTGACACCGCGAGGTGGAGTAAATCGCTATCTC 240
 Qy 241 GGAACATTAACAAGCGAGGTCTGGTAGCGACGATTCGTGGCGAAAAAGGGCTGCGCG 300
 Db 241 GGAACATTAACAAGCGAGGTCTGGTAGCGACGATTCGTGGCGAAAAAGGGCTGCGCG 300
 Qy 301 CAGTCTGATCAGGTGCGATTAATTCGTGATGTCAGACGATTAACCCGTAAACAGCGCA 360
 Db 301 CAGTCTGATCAGGTGCGATTAATTCGTGATGTCAGACGATTAACCCGTAAACAGCGCA 360
 Qy 361 GCTTTCGCGCGCGGGGTGGCACTGACCTGTCATTAACCTGATATCCAGGCAATACG 420
 Db 361 GCTTTCGCGCGCGGGGTGGCACTGACCTGTCATTAACCTGATATCCAGGCAATACG 420
 Qy 421 TCCAGTCCCGATCCGTTATTTAATCCTTAATAAAGTGGGTTTGGCACTGATTAACGG 480
 Db 421 TCCAGTCCCGATCCGTTATTTAATCCTTAATAAAGTGGGTTTGGCACTGATTAACGG 480
 Qy 481 AACGTGACTGACGCGATCTCTCAGCAGGAGGAGGTCAATTGCTGACTTTACCGGGCAT 540
 Db 481 AACGTGACTGACGCGATCTCTCAGCAGGAGGAGGTCAATTGCTGACTTTACCGGGCAT 540
 Qy 541 CGGCAAAAGCGGCTTTGCGCACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTTGTC 600
 Db 541 CGGCAAAAGCGGCTTTGCGCACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTTGTC 600
 Qy 601 CTTAAACGTGAAGAACAGGACGAAAGCTGTTCAATTAACGACGATTAACATCCGAACTTC 660
 Db 601 CTTAAACGTGAAGAACAGGACGAAAGCTGTTCAATTAACGACGATTAACATCCGAACTTC 660
 Qy 661 AAGGTGAGGCGCGACATATCTCATTAACGGGTGCGGTAAAGCTGATCAATGCTGACG 720
 Db 661 AAGGTGAGGCGCGACATATCTCATTAACGGGTGCGGTAAAGCTGATCAATGCTGACG 720
 Qy 721 GAGATATTTCTCTGCAACAGCAACAGGAAATGCGGAGCGGGGTGGGAAAGATCACC 780
 Db 721 GAGATATTTCTCTGCAACAGCAACAGGAAATGCGGAGCGGGGTGGGAAAGATCACC 780
 Qy 781 GATTACACCAAGTGAACAACCTTGTCAATTTGCAATTAAGCCCATTTTATTTGCTACAA 840
 Db 781 GATTACACCAAGTGAACAACCTTGTCAATTTGCAATTAAGCCCATTTTATTTGCTACAA 840
 Qy 841 CGCAGCGCAGAGGTGCGCGCAGCGCGCACCCCGTTATTTGATTTGATCAATGCGAGCG 900
 Db 841 CGCAGCGCAGAGGTGCGCGCAGCGCGCACCCCGTTATTTGATTTGATCAATGCGAGCG 900
 Qy 901 TTGACGCGCCATTCACCGCAAAAACAGCGGATGAGTGAATTAACCACTTCACTGATCTG 960
 Db 901 TTGACGCGCCATTCACCGCAAAAACAGCGGATGAGTGAATTAACCACTTCACTGATCTG 960
 Qy 961 TTTATTTGCGCGACAGATTAATTAATTCGTGGAGATTCGCGGCGCACTGAGAGTCAACTG 1020
 Db 961 TTTATTTGCGCGACAGATTAATTAATTCGTGGAGATTCGCGGCGCACTGAGAGTCAACTG 1020
 Qy 1021 ACGCTTCCCGGTCAAGCGCGGATTAACAGCGCGCAGAGGTGGAATCTGTGTTTAAAGCTG 1080
 Db 1021 ACGCTTCCCGGTCAAGCGCGGATTAACAGCGCGCAGAGGTGGAATCTGTGTTTAAAGCTG 1080
 Qy 1081 CGTGGGCTTAAGGATTAACAGCGAGTGAATTCAGGTTTCCGTGCTTCCAGACTTTACG 1140
 Db 1081 CGTGGGCTTAAGGATTAACAGCGAGTGAATTCAGGTTTCCGTGCTTCCAGACTTTACG 1140
 Qy 1141 CAGATGCGGTGAATAAAGCGCGCTGTCTTAATAATGCGCGCGCGAGAGGTGAATCTGAC 1200
 Db 1141 CAGATGCGGTGAATAAAGCGCGCTGTCTTAATAATGCGCGCGCGAGAGGTGAATCTGAC 1200
 Qy 1201 CTGGCAGATGTGAAGAGGAAATGCGCAGGCGATGTGTTGTTGCGAGGTTTACGCA 1260
 Db 1201 CTGGCAGATGTGAAGAGGAAATGCGCAGGCGATGTGTTGTTGCGAGGTTTACGCA 1260

Qy 1261 ATCGATGAAGCAACGATTAACCGCGGTGCAAGTTTGAATCTCATCAACCATCAATC 1320
 Db 1261 ATCGATGAAGCAACGATTAACCGCGGTGCAAGTTTGAATCTCATCAACCATCAATC 1320
 Qy 1321 TAA 1323
 Db 1321 TAA 1323
 RESULT 3
 US-10-034-985-1
 ; Sequence 1, Application US/10034985
 ; Publication No. US20030049815A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kretz, Keith
 ; TITLE OF INVENTION: NOVEL PHYTASE
 ; FILE REFERENCE: 09010/029003
 ; CURRENT APPLICATION NUMBER: US/10/034,985
 ; CURRENT FILING DATE: 2001-12-21
 ; PRIOR APPLICATION NUMBER: US/09/580,515
 ; PRIOR FILING DATE: 1999-05-25
 ; PRIOR APPLICATION NUMBER: 09/291,931
 ; PRIOR FILING DATE: 1999-04-13
 ; PRIOR APPLICATION NUMBER: 08/910,798
 ; PRIOR FILING DATE: 1997-08-13
 ; PRIOR APPLICATION NUMBER: 09/259,214
 ; PRIOR FILING DATE: 1999-03-01
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 1323
 ; TYPE: DNA
 ; ORGANISM: Escherichia coli
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(1320)
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(1323)
 ; OTHER INFORMATION: n = A,T,C or G
 US-10-034-985-1
 Query Match 100.0%; Score 1322; DB 6; Length 1323;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ATGAAAGGATCTTAATCCATTTTATCTCTGATTCGTTAAACCCCGAATCTGCA 60
 Db 1 ATGAAAGGATCTTAATCCATTTTATCTCTGATTCGTTAAACCCCGAATCTGCA 60
 Qy 61 TTGCGTCAAGTAGAGCCCGAGCTGAAAGTGAAGTGTGATTTGTCAGTCTGATGAT 120
 Db 61 TTGCGTCAAGTAGAGCCCGAGCTGAAAGTGAAGTGTGATTTGTCAGTCTGATGAT 120
 Qy 121 GTGCGTCTCAACCAAGCCGCAACGCAATGATGACAGATGTCACCCGAGCGCATGCGCA 180
 Db 121 GTGCGTCTCAACCAAGCCGCAACGCAATGATGACAGATGTCACCCGAGCGCATGCGCA 180
 Qy 181 ACGTGGCGGTAAACCTGGTTGGCTGACACCGCGAGGTGAGCTAATGCTTATCTC 240
 Db 181 ACGTGGCGGTAAACCTGGTTGGCTGACACCGCGAGGTGAGCTAATGCTTATCTC 240
 Qy 241 GGAACATTAACAAGCGAGGTCTGGTAGCGACGATTCGTGGCGAAAAAGGGCTGCGCG 300
 Db 241 GGAACATTAACAAGCGAGGTCTGGTAGCGACGATTCGTGGCGAAAAAGGGCTGCGCG 300
 Qy 301 CAGTCTGATCAGGTGCGATTAATTCGTGATGTCAGACGATTAACCCGTAAACAGCGCA 360
 Db 301 CAGTCTGATCAGGTGCGATTAATTCGTGATGTCAGACGATTAACCCGTAAACAGCGCA 360
 Qy 361 GCTTTCGCGCGCGGGGTGGCACTGACCTGTCATTAACCTGATATCCAGGCAATACG 420
 Db 361 GCTTTCGCGCGCGGGGTGGCACTGACCTGTCATTAACCTGATATCCAGGCAATACG 420

```
OY 421 TCCAGTCCCGATCCGTTATTAATCTCTAAAAATGCGGTTTGGCAATGATTAACCG 480
Db 421 TCCAGTCCCGATCCGTTATTAATCTCTAAAAATGCGGTTTGGCAATGATTAACCG 480
OY 481 AACGTGACTGACGCGATCCTGACGAGGAGGAGGTCATTTGCTGACTTTACCGGGCAT 540
Db 481 AACGTGACTGACGCGATCCTGACGAGGAGGAGGTCATTTGCTGACTTTACCGGGCAT 540
OY 541 CCGGAAACGGGCTTCCGGAATGGAACGGGTGCTTAATTTCCGATCAACTTGTGC 600
Db 541 CCGGAAACGGGCTTCCGGAATGGAACGGGTGCTTAATTTCCGATCAACTTGTGC 600
OY 601 CTTAAACGTGAAAAACAGAGCAAAAGCTGTTCACTTAACGAGGCTTACATCGGAATC 660
Db 601 CTTAAACGTGAAAAACAGAGCAAAAGCTGTTCACTTAACGAGGCTTACATCGGAATC 660
OY 661 AAGGTGACGCGCGACATGTCTCATTTAACCGGTGCGTAAGCTTCGATCAATGCTGACG 720
Db 661 AAGGTGACGCGCGACATGTCTCATTTAACCGGTGCGTAAGCTTCGATCAATGCTGACG 720
OY 721 GAGATTTTCTTCCTGCAACAAGCAAGGAAATGCGGAGCCGGGTGGGAAAGATCAC 780
Db 721 GAGATTTTCTTCCTGCAACAAGCAAGGAAATGCGGAGCCGGGTGGGAAAGATCAC 780
OY 781 GATTCAACACAGTGGAAACACTTGTGTAAGTTGCAATAACGCAATTTTATTTGCTACAA 840
Db 781 GATTCAACACAGTGGAAACACTTGTGTAAGTTGCAATAACGCAATTTTATTTGCTACAA 840
OY 841 CGCAGCGCAAGAGTTGCCCGAGCGCGCACCCCGTTATTTGATTTGATCAATGCGACG 900
Db 841 CGCAGCGCAAGAGTTGCCCGAGCGCGCACCCCGTTATTTGATTTGATCAATGCGACG 900
OY 901 TTGAGCCCCCATCCACCGGCAAAAAACAGGCGTATGTTGACATTAACCCACTTACTG 960
Db 901 TTGAGCCCCCATCCACCGGCAAAAAACAGGCGTATGTTGACATTAACCCACTTACTG 960
OY 961 TTTATTTGCGGACACGATTAATCTGCGAAATCTCGGCGGCGCACTGGAAGCTCAACTG 1020
Db 961 TTTATTTGCGGACACGATTAATCTGCGAAATCTCGGCGGCGCACTGGAAGCTCAACTG 1020
OY 1021 ACGCTTCCCGGTACGCGGATTAACAAGCGCGCAAGGTGTAATCTGTGTTTGAACGCTG 1080
Db 1021 ACGCTTCCCGGTACGCGGATTAACAAGCGCGCAAGGTGTAATCTGTGTTTGAACGCTG 1080
OY 1081 CGTGGCTTAAGCGATTAACGCGCATGTAAGTTCAAGTTTGGCTTTCACACTTTACAG 1140
Db 1081 CGTGGCTTAAGCGATTAACGCGCATGTAAGTTCAAGTTTGGCTTTCACACTTTACAG 1140
OY 1141 CAGATGCGTGAATTAACGCGCTGTCATTTAAATACGCGCGCGGAGAGGTGAACTGAC 1200
Db 1141 CAGATGCGTGAATTAACGCGCTGTCATTTAAATACGCGCGCGGAGAGGTGAACTGAC 1200
OY 1201 CTGGCAGGATGTAAGAGCGAAATGCGAGGCGCATGTGTTGTTGCGAGGTTTACGCA 1260
Db 1201 CTGGCAGGATGTAAGAGCGAAATGCGAGGCGCATGTGTTGTTGCGAGGTTTACGCA 1260
OY 1261 ATCGGTAATGAAGCAGCATACCGCGCTGCACTTGAATCTCATACATCAATCAAC 1320
Db 1261 ATCGGTAATGAAGCAGCATACCGCGCTGCACTTGAATCTCATACATCAATCAAC 1320
OY 1321 TAA 1323
Db 1321 TAA 1323
```

RESULT 4
US-10-430-356-1

; Sequence 1, Application US/10430356
; Publication No. US20030232041A1
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE

```
; FILE REFERENCE: 09010/029003  
; CURRENT APPLICATION NUMBER: US/10/430,356  
; PRIOR FILING DATE: 2003-05-05  
; PRIOR APPLICATION NUMBER: US/09/580,515  
; PRIOR FILING DATE: 1999-05-25  
; PRIOR APPLICATION NUMBER: 09/318,528  
; PRIOR FILING DATE: 1999-05-25  
; PRIOR APPLICATION NUMBER: 09/291,931  
; PRIOR FILING DATE: 1999-04-13  
; PRIOR APPLICATION NUMBER: 08/910,798  
; PRIOR FILING DATE: 1997-08-13  
; PRIOR APPLICATION NUMBER: 09/259,214  
; PRIOR FILING DATE: 1999-03-01  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1323  
; TYPE: DNA  
; ORGANISM: Escherichia coli  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(1320)  
; FEATURE:  
; NAME/KEY: misc_feature  
; LOCATION: (1)...(1323)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-430-356-1
```

```
Query Match 100.0%; Score 1322; DB 7; Length 1323;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 ATGAAAGGATCTTAATCCATTTTATCTCTGATTCGTTAACCCGCAATGCA 60  
Db 1 ATGAAAGGATCTTAATCCATTTTATCTCTGATTCGTTAACCCGCAATGCA 60  
OY 61 TTGCTCAGAGTGAAGCCGAGCTGAAGCTGAAAGTGTGATTTGTCAGTGTGATG 120  
Db 61 TTGCTCAGAGTGAAGCCGAGCTGAAGCTGAAAGTGTGATTTGTCAGTGTGATG 120  
OY 121 GTGCGTGTCCACCAAGGCGACGCAACTGATGAGATGTCACCCGACATGGCCA 180  
Db 121 GTGCGTGTCCACCAAGGCGACGCAACTGATGAGATGTCACCCGACATGGCCA 180  
OY 121 GTGCGTGTCCACCAAGGCGACGCAACTGATGAGATGTCACCCGACATGGCCA 180  
Db 121 GTGCGTGTCCACCAAGGCGACGCAACTGATGAGATGTCACCCGACATGGCCA 180  
OY 181 ACCTGGCGGTAATACGAGGTGCTGACACCGCGAGGTGAGTAAATCGCTATCTC 240  
Db 181 ACCTGGCGGTAATACGAGGTGCTGACACCGCGAGGTGAGTAAATCGCTATCTC 240  
OY 181 ACCTGGCGGTAATACGAGGTGCTGACACCGCGAGGTGAGTAAATCGCTATCTC 240  
Db 181 ACCTGGCGGTAATACGAGGTGCTGACACCGCGAGGTGAGTAAATCGCTATCTC 240  
OY 241 GGAATTACCAACGCGACGCTGTGTAGCGCATGCTGCGGAAAGGCTGCGCG 300  
Db 241 GGAATTACCAACGCGACGCGTGTGTAGCGCGAGTGTGCGGAAAGGCTGCGCG 300  
OY 301 CAGTGTGTCAGGTGCGATTAATGCTGATGTGACAGAGGTACCGTAAACAGGCGAA 360  
Db 301 CAGTGTGTCAGGTGCGATTAATGCTGATGTGACAGAGGTACCGTAAACAGGCGAA 360  
OY 361 GCTTTCGCGCGCGGCTGCGACCTGACGTGTGCAATTAACCGTACATCCAGGAGATACG 420  
Db 361 GCTTTCGCGCGCGGCTGCGACCTGACGTGTGCAATTAACCGTACATCCAGGAGATACG 420  
OY 421 TCCAGTCCCGATCCGTTATTAATCTCTAAAAATGCGGTTTGGCAATGATTAACCG 480  
Db 421 TCCAGTCCCGATCCGTTATTAATCTCTAAAAATGCGGTTTGGCAATGATTAACCG 480  
OY 481 AACGTGACTGACGCGATCCTGACGAGGAGGAGTCAATTTGCTGACTTTACCGGGCAT 540  
Db 481 AACGTGACTGACGCGATCCTGACGAGGAGGAGTCAATTTGCTGACTTTACCGGGCAT 540  
OY 541 CCGGAAACGGGCTTCCGGAATGGAACGGGTGCTTAATTTCCGATCAACTTGTGC 600  
Db 541 CCGGAAACGGGCTTCCGGAATGGAACGGGTGCTTAATTTCCGATCAACTTGTGC 600  
OY 601 CTTAAACGTGAAAAACAGAGCAAAAGCTGTTCACTTAACGAGGCTTACATCGGAATC 660
```

Db 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTACTTAAACGAGCATTAACATGGAACCTC 660
Qy 661 AAGGTAGAGCGCGACATATCTTATTACCGGTGGGTAAAGCTCCGATCATAGCTGAGC 720
Db 661 AAGGTAGAGCGCGACATATCTTATTACCGGTGGGTAAAGCTCCGATCATAGCTGAGC 720
Qy 721 GAGATATTTCTCTGCAACAGCAGAGGAAATGCGAGCCGGGGTGGGGAAGGATCACC 780
Db 721 GAGATATTTCTCTGCAACAGCAGAGGAAATGCGAGCCGGGGTGGGGAAGGATCACC 780
Qy 781 GATTCAACACAGTGAACAACCTTGTCTAAGTTTGCAATAGCGCAATTTATTGCTACAA 840
Db 781 GATTCAACACAGTGAACAACCTTGTCTAAGTTTGCAATAGCGCAATTTATTGCTACAA 840
Qy 841 CGCAGCGCCAGAGGTTGCGCGCAGCGCGCCACCCCGTTATGATTTGATCATGAGCG 900
Db 841 CGCAGCGCCAGAGGTTGCGCGCAGCGCGCCACCCCGTTATGATTTGATCATGAGCG 900
Qy 901 TTGAGCGCCCATCCACCGCAAAACAGGGGTATGATGATCACTTACCCCTTCAGTACTG 960
Db 901 TTGAGCGCCCATCCACCGCAAAACAGGGGTATGATGATCACTTACCCCTTCAGTACTG 960
Qy 961 TTTATTTGCCGACACAGATCTAATCTGCAAACTTCGGCGCGCACTGAGCTCAACTCG 1020
Db 961 TTTATTTGCCGACACAGATCTAATCTGCAAACTTCGGCGCGCACTGAGCTCAACTCG 1020
Qy 1021 ACCGTTCCCGGTCAAGCGGATACACGCGCCAGGTGTGAACCTGTGTTGAACGCTG 1080
Db 1021 ACCGTTCCCGGTCAAGCGGATACACGCGCCAGGTGTGAACCTGTGTTGAACGCTG 1080
Qy 1081 CGTCCGCTAAGGATTAACAGCCAGTGGATTTCAGGTTCCGTGCTTCCAGACTTACG 1140
Db 1081 CGTCCGCTAAGGATTAACAGCCAGTGGATTTCAGGTTCCGTGCTTCCAGACTTACG 1140
Qy 1141 CAGATCCGCTAATAACCGCGCTGTCTAATAATACGCGCCGAGAGGTGAACCTGACC 1200
Db 1141 CAGATCCGCTAATAACCGCGCTGTCTAATAATACGCGCCGAGAGGTGAACCTGACC 1200
Qy 1201 CTGGCAGATGTGAAGAGCGAAATGCGCAGGCGATGTGTTGTCAGGTTTAAACGAA 1260
Db 1201 CTGGCAGATGTGAAGAGCGAAATGCGCAGGCGATGTGTTGTCAGGTTTAAACGAA 1260
Qy 1261 ATCGGAATGAAGCAGCATACCGGCTGTCATTAATACGCGCCGAGAGGTGAACCTGAC 1320
Db 1261 ATCGGAATGAAGCAGCATACCGGCTGTCATTAATACGCGCCGAGAGGTGAACCTGAC 1320
Qy 1321 TAA 1323
Db 1321 TAA 1323

RESULT 5
US-10-601-319-1

Sequence 1, Application US/10601319
Publication No. US20040091968A1
GENERAL INFORMATION:
APPLICANT: Short, Jay M.
APPLICANT: Kretz, Keith A.
APPLICANT: Gray, Kevin A.
APPLICANT: Barton, Nelson Robert
APPLICANT: Garrett, James B.
APPLICANT: O'Donoghue, Eileen
APPLICANT: Mathur, Eric J.
TITLE OF INVENTION: RECOMBINANT PHYPHASES AND METHODS OF MAKING
FILE REFERENCE: 09010-029011
CURRENT APPLICATION NUMBER: US/10/601,319
CURRENT FILING DATE: 2003-06-20
PRIOR APPLICATION NUMBER: US 09/866,379
PRIOR FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: US 09/580,515
PRIOR FILING DATE: 2000-05-25

PRIOR APPLICATION NUMBER: US 09/318,528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: US 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 09/259,214
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: US 08/910,798
PRIOR FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1323
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1320)
FEATURE:
NAME/KEY: misc_feature
LOCATION: 216..feature
OTHER INFORMATION: n = A,T,C or G
US-10-601-319-1
Query Match 100.0%; Score 1322; DB 8; Length 1323;
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAAAGGATCTTAATCCATTTTATCTTCTGATTCGCTTAACCCGCAATCTGCA 60
Db 1 ATGAAAGGATCTTAATCCATTTTATCTTCTGATTCGCTTAACCCGCAATCTGCA 60
Qy 61 TTGCTCAGAGTGAACCGGAGCTGAAGCTGGAAGTGTGTGATTTGTCAGTGTATGCT 120
Db 61 TTGCTCAGAGTGAACCGGAGCTGAAGTGTGTGATTTGTCAGTGTATGCT 120
Qy 121 GTGCGTGTCCCAACCAAGGCGACGCAATGACAGATGCAACCCGCAAGGATGAGCA 180
Db 121 GTGCGTGTCCCAACCAAGGCGACGCAATGACAGATGCAACCCGCAAGGATGAGCA 180
Qy 181 ACCGCGCGGTAAACCTGGTTGCTGACACCGCGAGGTGTGAGCTAATCGCTATCTC 240
Db 181 ACCGCGCGGTAAACCTGGTTGCTGACACCGCGAGGTGTGAGCTAATCGCTATCTC 240
Qy 241 GGCATTTACCAACCGCAGCGTGTGTAGCCGAGATGCTGCGCAAAAGGGCTGCCG 300
Db 241 GGCATTTACCAACCGCAGCGTGTGTAGCCGAGATGCTGCGCAAAAGGGCTGCCG 300
Qy 301 CAGCTGATCAGGTGCGGATTTATGCTGATGTCGACGAGCGTAAACCGGTAAACAGCGAA 360
Db 301 CAGCTGATCAGGTGCGGATTTATGCTGATGTCGACGAGCGTAAACCGGTAAACAGCGAA 360
Qy 361 GCCTTCGCGCGCGGCTGCGACCTGACCTGCAATTAACCTGATACATACCGCAGATAGC 420
Db 361 GCCTTCGCGCGCGGCTGCGACCTGACCTGCAATTAACCTGATACATACCGCAGATAGC 420
Qy 421 TCAGTCCGATCCGTTATTTAATCTCTTAAAACTGCGCTTGGCAACTGAGTAAACGCG 480
Db 421 TCAGTCCGATCCGTTATTTAATCTCTTAAAACTGCGCTTGGCAACTGAGTAAACGCG 480
Qy 481 AACGTAAGTGAACCGGATCTGACGAGGCGAGAGGCTCAATTCGATCTTAACCGGGAT 540
Db 481 AACGTAAGTGAACCGGATCTGACGAGGCGAGAGGCTCAATTCGATCTTAACCGGGAT 540
Qy 541 CGGCAAAAGGCGTTTCGGAACCTGGAACGGGTGCTTAATTTTCGCAATCAACTTGTGC 600
Db 541 CGGCAAAAGGCGTTTCGGAACCTGGAACGGGTGCTTAATTTTCGCAATCAACTTGTGC 600
Qy 601 CTTAAACGTGAGAAACAGAGAAAGCTGTCTTAATACGAGGCTTAATCCATCGGAATC 660
Db 601 CTTAAACGTGAGAAACAGAGAAAGCTGTCTTAATACGAGGCTTAATCCATCGGAATC 660
Qy 661 AAGGTAGCGCGCAATATGCTCATTAACCGGTGGGTGAGCTTCGATCATGCTGAGC 720

Db 661 AAGTGAGCGCGCAATGTCTATTAAACGGTGGGTAAGCTCGCATCAATGCTGACG 720
Qy 721 GAGATATTTCTCTGCAACAAGCAGAGGAATGCCGAGCCGGGGTGGGAAAGATCAC 780
Db 721 GAGATATTTCTCTGCAACAAGCAGAGGAATGCCGAGCCGGGGTGGGAAAGATCAC 780
Qy 781 GATTACACGAGTGGAAACCTGTAGTTTGTGATAAGCGGCAATTTTATTTGCTACAA 840
Db 781 GATTACACGAGTGGAAACCTGTAGTTTGTGATAAGCGGCAATTTTATTTGCTACAA 840
Qy 841 CGCAGCGCAGAGTGTCCCGCAGCCGCGCAACCCGTTATTTGATTTGATGATGCAAGC 900
Db 841 CGCAGCGCAGAGTGTCCCGCAGCCGCGCAACCCGTTATTTGATTTGATGATGCAAGC 900
Qy 901 TTGACGCCCATCCACCGGCAAAAACAGCGCGTATGTTGACATTAACCACTTCAGTACTG 960
Db 901 TTGACGCCCATCCACCGGCAAAAACAGCGCGTATGTTGACATTAACCACTTCAGTACTG 960
Qy 961 TTATATGCGCGACACGATACTAATCTGGCAAAATCTCGGCGGCGCATGGAGCTCAACTGG 1020
Db 961 TTATATGCGCGACACGATACTAATCTGGCAAAATCTCGGCGGCGCATGGAGCTCAACTGG 1020
Qy 1021 ACGCTTCCCGGTCAAGCCGGAATACACGCGCCGCAAGGTGTGAATGTTGTAACGCTGG 1080
Db 1021 ACGCTTCCCGGTCAAGCCGGAATACACGCGCCGCAAGGTGTGAATGTTGTAACGCTGG 1080
Qy 1081 CGTGGCGTAAACGATTAACAGCAGTGTGATTCAGGTTTGGCTGTTTCAACATTTACAG 1140
Db 1081 CGTGGCGTAAACGATTAACAGCAGTGTGATTCAGGTTTGGCTGTTTCAACATTTACAG 1140
Qy 1141 CAGATGCGTAAACGCGCTGTCTATTAATACGCGCCGCGGAGAGGTGAAACGTCAGC 1200
Db 1141 CAGATGCGTAAACGCGCTGTCTATTAATACGCGCCGCGGAGAGGTGAAACGTCAGC 1200
Qy 1201 CTGGCAGGATGTGAAGACGAATGCGCAGGCGCATGTGTTGTTGCGAGGTTTACGCA 1260
Db 1201 CTGGCAGGATGTGAAGACGAATGCGCAGGCGCATGTGTTGTTGCGAGGTTTACGCA 1260
Qy 1261 ATCGTGAATGAAGCAGCATACCGGCGTGCAGTTTGAATCTCATCACATCACATAC 1320
Db 1261 ATCGTGAATGAAGCAGCATACCGGCGTGCAGTTTGAATCTCATCACATCACATAC 1320
Qy 1321 TAA 1323
Db 1321 TAA 1323

RESULT 6

US-10-933-115-1
; Sequence 1, Application US/10933115
; Publication No. US20050281792A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garret, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES
; FILE REFERENCE: 564462001822
; CURRENT APPLICATION NUMBER: US/10/933,115
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214

;; PRIOR FILING DATE: 1999-03-01
;; PRIOR APPLICATION NUMBER: US 08/910,798
;; PRIOR FILING DATE: 1997-08-13
;; NUMBER OF SEQ ID NOS: 10
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 1
;; LENGTH: 1323
;; TYPE: DNA
;; ORGANISM: Escherichia coli
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (1)...(1320)
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: 216
;; OTHER INFORMATION: n = A,T,C or G
US-10-933-115-1
Query Match 100.0%; Score 1322; DB 10; Length 1323;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAAGCGATCTTAATCCATTTTATCTCTCTGATTCGTTAACCCGCAATCTGCA 60
Db 1 ATGAAGCGATCTTAATCCATTTTATCTCTCTGATTCGTTAACCCGCAATCTGCA 60
Qy 61 TTGCTCAGAGTGAAGCCGAGCTGAAGCTGGAAGTGTGATTTGATCTGATGCT 120
Db 61 TTGCTCAGAGTGAAGCCGAGCTGAAGCTGGAAGTGTGATTTGATCTGATGCT 120
Qy 121 GTGAGTCTCCAAACCAAGGCGCAGCACTGATGAGGATGACCCGAGCGCATGGCCA 180
Db 121 GTGAGTCTCCAAACCAAGGCGCAGCACTGATGAGGATGACCCGAGCGCATGGCCA 180
Qy 181 ACCGCGCGGTAACCTGAGTGTGCTGACACCGCGAGTGTGAGTATCGCTATCTC 240
Db 181 ACCGCGCGGTAACCTGAGTGTGCTGACACCGCGAGTGTGAGTATCGCTATCTC 240
Qy 241 GGAATTAACCAAGCGCAGCGCTGTGTAGCCGAGCGAATGCTGCGGAAAGGCTGCCG 300
Db 241 GGAATTAACCAAGCGCAGCGCTGTGTAGCCGAGCGAATGCTGCGGAAAGGCTGCCG 300
Qy 301 CAGCTGTGAGGTCGATTAATTTGCTGATGTGAGAGGCTGCTAAACAGGCGAA 360
Db 301 CAGCTGTGAGGTCGATTAATTTGCTGATGTGAGAGGCTGCTAAACAGGCGAA 360
Qy 361 GCCTTCCGCGCGGCTGAGCACTGATGTGACATTAACCGTACATACCGAGGAGATACG 420
Db 361 GCCTTCCGCGCGGCTGAGCACTGATGTGACATTAACCGTACATACCGAGGAGATACG 420
Qy 421 TCCAGTCCGATCCGTTATTTATCTCTTAATAACCTGCGTTTCCCACTGATACGCG 480
Db 421 TCCAGTCCGATCCGTTATTTATCTCTTAATAACCTGCGTTTCCCACTGATACGCG 480
Qy 481 AACGTGACAGCGGATCTCAGAGGCGAGGAGGTCATTTGCTTAACCGGCGCAT 540
Db 481 AACGTGACAGCGGATCTCAGAGGCGAGGAGGTCATTTGCTTAACCGGCGCAT 540
Qy 541 CGGCAACGCGGTTCCGCACTGGAACGGGTCCTTAATTTCCGCAATCAACTGTGC 600
Db 541 CGGCAACGCGGTTCCGCACTGGAACGGGTCCTTAATTTCCGCAATCAACTGTGC 600
Qy 601 CTTAAAGTGAAGAAACAGAGCAAGAGCTGTTCAATTAACGAGGATTAACATCGGAATC 660
Db 601 CTTAAAGTGAAGAAACAGAGCAAGAGCTGTTCAATTAACGAGGATTAACATCGGAATC 660
Qy 661 AAGGTGAGCGCGCAATGTCTCATTAACCGGTGCGGTAAAGCTCGATCAATGCTGACG 720
Db 661 AAGGTGAGCGCGCAATGTCTCATTAACCGGTGCGGTAAAGCTCGATCAATGCTGACG 720
Qy 721 GAGATATTTCTCTGCAACAAGCAGAGGATGCGGAGCGGGGTTGGGAAAGATCAC 780
Db 721 GAGATATTTCTCTGCAACAAGCAGAGGATGCGGAGCGGGGTTGGGAAAGATCAC 780

QY 781 GATTACACAGAGGAAACCTTGTAGTTGGCATAGCGCAATTTTATTGCTACAA 840
DB 781 GATTACACAGAGGAAACCTTGTAGTTGGCATAGCGCAATTTTATTGCTACAA 840
QY 841 CGCAGCCGAGAGGTTGCCCGCAGCCCGCAACCCCGTTATTGATTTGATGAGGAGC 900
DB 841 CGCAGCCGAGAGGTTGCCCGCAGCCCGCAACCCCGTTATTGATTTGATGAGGAGC 900
QY 901 TTGACGCCCCCATTCACCGCAAAAACGGGGTATGTGTGACATTAACCACTTCAGTACG 960
DB 901 TTGACGCCCCCATTCACCGCAAAAACGGGGTATGTGTGACATTAACCACTTCAGTACG 960
QY 961 TTATATGCGGAGACAGATATCTATCTGGAATCTGGGGGGGCACTGGAGCTCAACTGG 1020
DB 961 TTATATGCGGAGACAGATATCTGGAATCTGGGGGGGCACTGGAGCTCAACTGG 1020
QY 1021 AGCGTTCCCGGTACGCGGATTAACAGCGCGCAGGTGTGAATCTGGTTTGAACGCTGG 1080
DB 1021 AGCGTTCCCGGTACGCGGATTAACAGCGCGCAGGTGTGAATCTGGTTTGAACGCTGG 1080
QY 1081 CGTGGCTAAGCGATTAACAGCGAGTGAATTCAGTTTCTGCTTCCAGACTTTACAG 1140
DB 1081 CGTGGCTAAGCGATTAACAGCGAGTGAATTCAGTTTCTGCTTCCAGACTTTACAG 1140
QY 1141 CAGATGCGGTATTAACCGCGCTGTATTAATACCGCGCGGAGAGGTGAACCTGACC 1200
DB 1141 CAGATGCGGTATTAACCGCGCTGTATTAATACCGCGCGGAGAGGTGAACCTGACC 1200
QY 1201 CTGGCAGATGTAAGAGGAAATGCGCAGGGCATGTGTTCTGTTGACAGTTTACGAA 1260
DB 1201 CTGGCAGATGTAAGAGGAAATGCGCAGGGCATGTGTTCTGTTGACAGTTTACGAA 1260
QY 1261 ATCGTAATGAAGCAGCATACCGGGTSCAGTTTGAATCTCATCAATCAACATCAAC 1320
DB 1261 ATCGTAATGAAGCAGCATACCGGGTSCAGTTTGAATCTCATCAATCAACATCAAC 1320
QY 1321 TAA 1323
DB 1321 TAA 1323

RESULT 7
US-09-866-379-7
; Sequence 7, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETT, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 7
; LENGTH: 1901
; TYPE: DNA

ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-7
Query Match 97.4%; Score 1287.8; DB 3; Length 1901;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1293; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 ATGAAGGAGATCTTAATCCATTTTATCTCTTGAATTCGTTAACCCCGCAATCTGCA 60
DB 188 ATGAAGGAGATCTTAATCCATTTTATCTCTTGAATTCGTTAACCCCGCAATCTGCA 247
QY 61 TTGCGTCAAGAGTGAAGCCGAGAGTGAAGCTGGAAGTGTGTGATTTGATGTCATGCT 120
DB 248 TTGCGTCAAGAGTGAAGCCGAGAGTGAAGCTGGAAGTGTGTGATTTGATGTCATGCT 307
QY 121 GTGCGGTCTCCAAACCAAGGCCAGCACTGAATGAGATGTCACCCAGAGCATGAGCA 180
DB 308 GTGCGGTCTCCAAACCAAGGCCAGCACTGAATGAGATGTCACCCAGAGCATGAGCA 367
QY 181 ACCTGGCCGGTAAACTGGGTTGGCTGACACCGCGAGGTGTGAGTGAATCGCTATCTC 240
DB 368 ACCTGGCCGGTAAACTGGGTTGGCTGACACCGCGAGGTGTGAGTGAATCGCTATCTC 427
QY 241 GGCATTTACCAACGCGCAGCGTCTGTAGCCGACGAGATTGCTGGCGCAAAAAGGCGTCCG 300
DB 428 GGCATTTACCAACGCGCAGCGTCTGTAGCCGACGAGATTGCTGGCGCAAAAAGGCGTCCG 487
QY 301 CAGTCTGGTCAAGTGGCGATTTTGTCTGATGTGACAGACCTTAATCCCGTAAACAGGCGAA 360
DB 488 CAGTCTGGTCAAGTGGCGATTTTGTCTGATGTGACAGACCTTAATCCCGTAAACAGGCGAA 547
QY 361 GCCTTCGCGCGCGGCTGGACCTGATGTCATTAACGTTACATTAACCGAGCAGATACG 420
DB 548 GCCTTCGCGCGCGGCTGGACCTGATGTCATTAACGTTACATTAACCGAGCAGATACG 607
QY 421 TCCAGTCCCGATCCGTTATTTAATCTCTAATAAATCGGCGTTTGCCTGATTAACGCG 480
DB 608 TCCAGTCCCGATCCGTTATTTAATCTCTAATAAATCGGCGTTTGCCTGATTAACGCG 667
QY 481 AACGTGATGACGCGATCTCAGCAGGCGAGAGGCTCAATTGCTGACCTTAAACCGGCGAT 540
DB 668 AACGTGATGACGCGATCTCAGCAGGCGAGAGGCTCAATTGCTGACCTTAAACCGGCGAT 727
QY 541 CGGCAAAAGGGGTTTCGGAAGTGAAGGGGCTTAATTTTCGCGATCAAACTTGTGC 600
DB 728 CGGCAAAAGGGGTTTCGGAAGTGAAGGGGCTTAATTTTCGCGATCAAACTTGTGC 787
QY 601 CTTAAACGTGAGAAACAGAGCAAAAAGCTTTCAATTAACGAGCAGCTTAACATCGAATC 660
DB 788 CTTAAACGTGAGAAACAGAGCAAAAAGCTTTCAATTAACGAGCAGCTTAACATCGAATC 847
QY 661 AAGGTGAGCGCTGACATGTCTCAATTAACCGGTGGTGAAGCTTGCATCAATGCTGAGC 720
DB 848 AAGGTGAGCGCTGACATGTCTCAATTAACCGGTGGTGAAGCTTGCATCAATGCTGAGC 907
QY 721 GAGATATTTCTCTGCAACAGCAAGGAAATGCGGAGCGGGGTGGGAAAGATCC 780
DB 908 GAGATATTTCTCTGCAACAGCAAGGAAATGCGGAGCGGGGTGGGAAAGATCC 967
QY 781 GATTACACAGTGAACACCTTGTAGTTGCAATAGCGCAATTTTATTGCTACAA 840
DB 968 GATTACACAGTGAACACCTTGTAGTTGCAATAGCGCAATTTTATTGCTACAA 1027
QY 841 CGCAGCCGAGAGTGGCCCGCAGCCGCGCAACCCCGTTATTGATTTGATGAGCAGC 900
DB 1028 CGCAGCCGAGAGTGGCCCGCAGCCGCGCAACCCCGTTATTGATTTGATGAGCAGC 1087
QY 901 TTGACGCCCCCATTCACCGCAAAAACGGGGTATGTGTGACATTAACCACTTCAGTACG 960

Db 1088 TTGACGCCCATCCACCCGAAACAGCGGTATGTTGATTAATCCCACTTCAGTGTCTG 1147
 Qy 961 TTATATGCGGACACAGATTAATCTGGAATATCGGCGGCGCATCTGAGCTCAACTGG 1020
 Db 1148 TTATATGCGGACACAGATTAATCTGGAATATCGGCGGCGCATCTGAGCTCAACTGG 1207
 Qy 1021 ACGCTTCCCGGTACGCCGATTAACAGCGCGCGAGGTGTGAATCTGATTTGAACGCTGG 1080
 Db 1208 ACGCTTCCCGGTACGCCGATTAACAGCGCGCGAGGTGTGAATCTGATTTGAACGCTGG 1267
 Qy 1081 CGTGGCTTAAGCATTAACAGCCAGTGTATTCAGTTTCGCTGCTTTCAGACTTTACAG 1140
 Db 1268 CGTGGCTTAAGCATTAACAGCCAGTGTATTCAGTTTCGCTGCTTTCAGACTTTACAG 1327
 Qy 1141 CAGATGCTGATTAACAGCCGCTGTCTTAATATAGCCGCCGAGAGGTGAACCTGACC 1200
 Db 1328 CAGATGCTGATTAACAGCCGCTGTCTTAATATAGCCGCCGAGAGGTGAACCTGACC 1387
 Qy 1201 CTGGCAGATGTGAAGACGAAATCGCAGGCGATGTCTGTTGGCAGTTTTCAGCAA 1260
 Db 1388 CTGGCAGATGTGAAGACGAAATCGCAGGCGATGTCTGTTGGCAGTTTTCAGCAA 1447
 Qy 1261 ATCGTGAATGAACGACGATTAACCGGCGGTGCGAGTTTGAGAT 1300
 Db 1448 ATCGTGAATGAACGACGATTAACCGGCGGTGCGAGTTTGAT 1487

RESULT 8
 US-10-156-660-3
 / Sequence 3, Application US/10156660
 / Publication No. US20030103958A1
 / GENERAL INFORMATION:
 / APPLICANT: Short, Jay M.
 / APPLICANT: Kretz, Keith
 / APPLICANT: Gray, Kevin A.
 / APPLICANT: Barton, Nelson R.
 / APPLICANT: Garrett, James B.
 / APPLICANT: O'Donoghue, Eileen
 / APPLICANT: Mathur, Eric J.
 / TITLE OF INVENTION: PHYLASES, NUCLEIC ACIDS ENCODING THEM
 / TITLE OF INVENTION: AND METHODS FOR MAKING AND USING THEM
 / FILE REFERENCE: 09010-029007
 / CURRENT APPLICATION NUMBER: US/10/156,660
 / PRIOR FILING DATE: 2002-10-01
 / PRIOR APPLICATION NUMBER: US 09/866,379
 / PRIOR FILING DATE: 2001-05-24
 / PRIOR APPLICATION NUMBER: US 09/580,515
 / PRIOR FILING DATE: 2000-05-25
 / PRIOR APPLICATION NUMBER: US 09/318,528
 / PRIOR FILING DATE: 1999-05-25
 / PRIOR APPLICATION NUMBER: US 09/291,931
 / PRIOR FILING DATE: 1999-04-13
 / PRIOR APPLICATION NUMBER: US 09/259,214
 / PRIOR FILING DATE: 1999-03-01
 / PRIOR APPLICATION NUMBER: US 08/910,798
 / PRIOR FILING DATE: 1997-08-13
 / NUMBER OF SEQ ID NOS: 4
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO 3
 / LENGTH: 1901
 / TYPE: DNA
 / ORGANISM: Escherichia coli
 / FEATURE:
 / NAME/KEY: CDS
 / LOCATION: (188) ... (1483)
 / FEATURE:
 / NAME/KEY: misc_feature
 / LOCATION: 403
 / OTHER INFORMATION: n = A,T,C or G
 US-10-156-660-3

Query Match 97.4%; Score 1287.8; DB 6; Length 1901;
 Best Local Similarity 99.5%; Pred. No. 0;

Matches 1293; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 Qy 1 ATGAAAGCATTTAATCCATTTTATCTTCTGATTCGCTTAACCCGCAATCTGCA 60
 Db 188 ATGAAAGCATTTAATCCATTTTATCTTCTGATTCGCTTAACCCGCAATCTGCA 247
 Qy 61 TTCCTCAGAGTGAAGCCGAGCTGAACCTGAAAGTGTGATTCAGTTCATAGT 120
 Db 248 TTCCTCAGAGTGAAGCCGAGCTGAACCTGAAAGTGTGATTCAGTTCATAGT 307
 Qy 121 GTGCTGCTCCACCAAGGCCAGCACTGATGCAAGTGTACCCAGACGATGGCCA 180
 Db 308 GTGCTGCTCCACCAAGGCCAGCACTGATGCAAGTGTACCCAGACGATGGCCA 367
 Qy 181 ACCGCGCGGTAACCTGGGTGGCTGACACCGGAGGTGGTGAATTCGCTATCTC 240
 Db 368 ACCGCGCGGTAACCTGGGTGGCTGACACCGGAGGTGGTGAATTCGCTATCTC 427
 Qy 241 GGAATTAAACCAACGACCGCTGTGTAGCCGACGATGCTGGCGAAGAGGCTGCCG 300
 Db 428 GGAATTAAACCAACGACCGCTGTGTAGCCGACGATGCTGGCGAAGAGGCTGCCG 487
 Qy 301 CAGTCTGCTCAAGTTCGCGATTAATGCTGATGTGACGACGATACCGCTAAACAGCGAA 360
 Db 488 CAGTCTGCTCAAGTTCGCGATTAATGCTGATGTGACGACGATACCGCTAAACAGCGAA 547
 Qy 361 GCCCTGCGCGCGGCGTGGGACCTGACTGTGCAATACCGTACATCCGAGCGATACG 420
 Db 548 GCCCTGCGCGCGGCGTGGGACCTGACTGTGCAATACCGTACATCCGAGCGATACG 607
 Qy 421 TCCAGTCCCGATCCGTTATTTAATCTCTAATAACCTGCGCTTGGCCAACTGATACGCG 480
 Db 608 TCCAGTCCCGATCCGTTATTTAATCTCTAATAACCTGCGCTTGGCCAACTGATACGCG 667
 Qy 481 AACGTACTGAAGCGATCTCAGACGAGCGCAGAGGGTCAATTGCTGACTTTACCGGCAT 540
 Db 668 AACGTACTGAAGCGATCTCAGACGAGCGCAGAGGGTCAATTGCTGACTTTACCGGCAT 727
 Qy 541 CGGCAACGGCGTTTGGCGAAGTGGAAAGGGGTCTTAATTTCCGCAATCAAACTTGTGC 600
 Db 728 CGGCAACGGCGTTTGGCGAAGTGGAAAGGGGTCTTAATTTCCGCAATCAAACTTGTGC 787
 Qy 601 CTTAAAGTGAAGAAACAGACGAAAGCTGTTCAATTAACGACGATTAACATCGAACTC 660
 Db 788 CTTAAAGTGAAGAAACAGACGAAAGCTGTTCAATTAACGACGATTAACATCGAACTC 847
 Qy 661 AAGGTAGCGCGCAATGTCTCATTAACCGGTGCGGTAAAGCTCGCATCAATGCTGACG 720
 Db 848 AAGGTAGCGCGCAATGTCTCATTAACCGGTGCGGTAAAGCTCGCATCAATGCTGACG 907
 Qy 721 GAGATATTTCTCTCGCAACAGCAACAGGAATGCCGAGCCGGGTGGGAAGATCAAC 780
 Db 908 GAGATATTTCTCTCGCAACAGCAACAGGAATGCCGAGCCGGGTGGGAAGATCAAC 967
 Qy 781 GATTCAACCAAGTGAACACCTTGTCTAAGTTGTGATTAACGCGCAATTTATTTGCTCAAA 840
 Db 968 GATTCAACCAAGTGAACACCTTGTCTAAGTTGTGATTAACGCGCAATTTATTTGCTCAAA 1027
 Qy 841 CGCAGCGCAGAGGTTCGCGGACCGCGCAACCCCGTTATTTGAATTTGATCATGAGCG 900
 Db 1028 CGCAGCGCAGAGGTTCGCGGACCGCGCAACCCCGTTATTTGAATTTGATCATGAGCG 1087
 Qy 901 TTGACGCCCATTCACCGCAAAAAACAGCGGTATGTGTGACATTAACCATTCGATGACTG 960
 Db 1088 TTGACGCCCATTCACCGCAAAAAACAGCGGTATGTGTGACATTAACCATTCGATGACTG 1147
 Qy 961 TTATATGCGGACACAGATTAATCTGGAATATCGGCGGCGCATCTGAGCTCAACTGG 1020
 Db 1148 TTATATGCGGACACAGATTAATCTGGAATATCGGCGGCGCATCTGAGCTCAACTGG 1207
 Qy 1021 ACGCTTCCCGGTACGCCGATTAACAGCGCGCGAGGTGTGAATCTGATTTGAACGCTGG 1080
 Db 1208 ACGCTTCCCGGTACGCCGATTAACAGCGCGCGAGGTGTGAATCTGATTTGAACGCTGG 1267

QY 1081 CGTGGCTAGAGGATTAACGCGCAGTGGATTGCGGTCTTCCAGACTTTACAG 1140
 Db 1268 CGTGGCTAGAGGATTAACGCGCAGTGGATTGCGGTCTTCCAGACTTTACAG 1327
 QY 1141 CAGATCGGTATTAACGCGCAGTGGATTGCGGTCTTCCAGACTTTACAG 1200
 Db 1328 CAGATCGGTATTAACGCGCAGTGGATTGCGGTCTTCCAGACTTTACAG 1387
 QY 1201 CTGGCAGATGTAAGAGCGAAATGCGCAGGAGCATGTGTTGTTGCAAGTTTACGCA 1260
 Db 1388 CTGGCAGATGTAAGAGCGAAATGCGCAGGAGCATGTGTTGTTGCAAGTTTACGCA 1447
 QY 1261 ATCGTAATGAAGCAGCATACCGGCGTGCAGTTTGAAT 1300
 Db 1448 ATCGTAATGAAGCAGCATACCGGCGTGCAGTTTGAAT 1487
 RESULT 9
 US-10-601-319-7
 / Sequence 7, Application US/10601319
 / Publication No. US20040091968A1
 / GENERAL INFORMATION:
 / APPLICANT: Short, Jay M.
 / APPLICANT: Kreitz, Keith A.
 / APPLICANT: Gray, Kevin A.
 / APPLICANT: Barton, Nelson Robert
 / APPLICANT: Garrett, James B.
 / APPLICANT: O'Donoghue, Eileen
 / APPLICANT: Mathur, Eric J.
 / TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
 / FILE REFERENCE: 09010-029011
 / CURRENT APPLICATION NUMBER: US/10/601,319
 / CURRENT FILING DATE: 2003-06-20
 / PRIOR APPLICATION NUMBER: US 09/866,379
 / PRIOR FILING DATE: 2001-05-24
 / PRIOR APPLICATION NUMBER: US 09/580,515
 / PRIOR FILING DATE: 2000-05-25
 / PRIOR APPLICATION NUMBER: US 09/318,528
 / PRIOR FILING DATE: 1999-05-25
 / PRIOR APPLICATION NUMBER: US 09/291,931
 / PRIOR FILING DATE: 1999-04-13
 / PRIOR APPLICATION NUMBER: US 09/259,214
 / PRIOR FILING DATE: 1999-03-01
 / PRIOR APPLICATION NUMBER: US 08/910,798
 / PRIOR FILING DATE: 1997-08-13
 / NUMBER OF SEQ ID NOS: 10
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO 7
 / LENGTH: 1901
 / TYPE: DNA
 / ORGANISM: Escherichia coli
 / FEATURE:
 / NAME/KEY: CDS
 / LOCATION: (188)...(1483)
 / FEATURE:
 / NAME/KEY: misc_feature
 / LOCATION: 403
 / OTHER INFORMATION: n = A,T,C or G
 / US-10-601-319-7
 Query Match 97.4%; Score 1287.8; DB 8; Length 1901;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1933; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 121 GTGCTGCTTCAACCAAGGCAAGCACTGATGCAAGATGTCAACCCAGACGATGCGCA 180
 Db 308 GTGCTGCTTCAACCAAGGCAAGCACTGATGCAAGATGTCAACCCAGACGATGCGCA 367
 QY 181 ACCGCGCGGTAAACGCGGTGCGTGAACCGCGGTGCGTGAAGTAAATGCGTAAATC 240
 Db 368 ACCGCGCGGTAAACGCGGTGCGTGAACCGCGGTGCGTGAAGTAAATGCGTAAATC 427
 QY 241 GGAATTATCAACGCGGCGTGTGAGCGGAGATTTGCGGCGAATAAGGCGTCCCG 300
 Db 428 GGAATTATCAACGCGGCGTGTGAGCGGAGATTTGCGGCGAATAAGGCGTCCCG 487
 QY 301 CAGTGTGCTGAGGTGCGGATTTATGCTGATGTGACGAGCGTACCCGTAACAGCGCA 360
 Db 488 CAGTGTGCTGAGGTGCGGATTTATGCTGATGTGACGAGCGTACCCGTAACAGCGCA 547
 QY 361 GCCTTGCGCGCGGCGTGCACCTGACGTGCAATTAACGCTATTAACGAGCGATAG 420
 Db 548 GCCTTGCGCGCGGCGTGCACCTGACGTGCAATTAACGCTATTAACGAGCGATAG 607
 QY 421 TCCAGTCCCGATCCGTTATTTATATCTTAAACCTGAGCTTGCACCTGATTAACGG 480
 Db 608 TCCAGTCCCGATCCGTTATTTATATCTTAAACCTGAGCTTGCACCTGATTAACGG 667
 QY 481 AACGTGATGACGCGGATCTTCAAGCGGCAAGAGGCTCAATGCTGACTTTACCGGCA 540
 Db 668 AACGTGATGACGCGGATCTTCAAGCGGCAAGAGGCTCAATGCTGACTTTACCGGCA 727
 QY 541 CGGCAAAAGGGGTTTCGGCACTGGAAGGGGCTTAATTTCCGAACTCAACCTTGCG 600
 Db 728 CGGCAAAAGGGGTTTCGGCACTGGAAGGGGCTTAATTTTCGAACTCAACCTTGCG 787
 QY 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGACGATTAACATCGAATC 660
 Db 788 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGACGATTAACATCGAATC 847
 QY 661 AAGGTGAGCGCGCAATGTCTCAATTAACCGGTGCGTGAAGCTTGCATCAATGCTGAC 720
 Db 848 AAGGTGAGCGCGCAATGTCTCAATTAACCGGTGCGTGAAGCTTGCATCAATGCTGAC 907
 QY 721 GAGATATTTCTCTGCAACAGCAAGGAAAGCGGAGCGGCGGCGGAGAGATCC 780
 Db 908 GAGATATTTCTCTGCAACAGCAAGGAAAGCGGAGCGGCGGCGGAGAGATCC 967
 QY 781 GATTCAACAGTGAACACTTGTGAATTTGCAATACGCGCAATTTTATTTGCTACAA 840
 Db 968 GATTCAACAGTGAACACTTGTGAATTTGCAATACGCGCAATTTTATTTGCTACAA 1027
 QY 841 CGCAGCGCAAGGTTGCGCGCAGCGCGCACCCCGTTATGATTTGATTCATGCGACG 900
 Db 1028 CGCAGCGCAAGGTTGCGCGCAGCGCGCACCCCGTTATGATTTGATTCATGCGACG 1087
 QY 901 TTGAGCGCCCATCCACCGCAAAAACAGCGTATGCTGACATTAACCACTTCAATACG 960
 Db 1088 TTGAGCGCCCATCCACCGCAAAAACAGCGTATGCTGACATTAACCACTTCAATACG 1147
 QY 961 TTTATTTGCGGACAGATATTAATCTGCAATCTGCGCGGCGCACTGAGGCTCAATCG 1020
 Db 1148 TTTATTTGCGGACAGATATTAATCTGCAATCTGCGCGGCGCACTGAGGCTCAATCG 1207
 QY 1021 ACGCTTCCCGGTCAAGCGGATTAACGCGCGCGAGGTGTAATCTGTTTGAACGCTGG 1080
 Db 1208 ACGCTTCCCGGTCAAGCGGATTAACGCGCGCGAGGTGTAATCTGTTTGAACGCTGG 1267
 QY 1081 CGTGGCTAAGGATTAACGCGAGGATTAAGGTTGCGTGTTCAGACTTTACG 1140
 Db 1268 CGTGGCTAAGGATTAACGCGAGGATTAAGGTTGCGTGTTCAGACTTTACG 1327
 QY 1141 CAGATCGGTATTAACGCGCAGTGGATTGCGGTCTTCCAGACTTTACAG 1200
 Db 1328 CAGATCGGTATTAACGCGCAGTGGATTGCGGTCTTCCAGACTTTACAG 1387

OY 1201 CTGCGAGGATGTGAAGAGCGAAATGCGAGGGGCAATGTGCTGGAGGTTTAAACGCA 1260
DB 1388 CTGGCAGGATGTGAAGAGCGAAATGCGAGGGGCAATGTGCTGGAGGTTTAAACGCA 1447
OY 1261 ATCGTAATGAAGCAGCATACCGGCGTGCAGTTTGAGAT 1300
DB 1448 ATCGTAATGAAGCAGCATACCGGCGTGCAGTTTGATAT 1487

RESULT 10
US-10-933-115-7

Sequence 7, Application US/10933115
Publication No. US20050281792A1

GENERAL INFORMATION:

APPLICANT: Short, Jay M.

APPLICANT: Kretz, Keith A.

APPLICANT: Gray, Kevin A.

APPLICANT: Barton, Nelson Robert

APPLICANT: Garrett, James B.

APPLICANT: O' Donoghue, Eileen

APPLICANT: Machur, Eric J.

TITLE OF INVENTION: RECOMBINANT BACTERIAL PHOSPHATASES AND USES

TITLE OF INVENTION: THEREOF

FILE REFERENCE: 56462001822

CURRENT FILING DATE: US/10/933,115

PRIOR FILING DATE: 2004-09-01

PRIOR FILING DATE: US/09/866,379

PRIOR FILING DATE: 2001-05-24

PRIOR APPLICATION NUMBER: US 09/580,515

PRIOR FILING DATE: 2000-05-25

PRIOR APPLICATION NUMBER: US 09/318,528

PRIOR FILING DATE: 1999-05-25

PRIOR APPLICATION NUMBER: US 09/291,931

PRIOR FILING DATE: 1999-04-13

PRIOR APPLICATION NUMBER: US 09/259,214

PRIOR FILING DATE: 1999-03-01

PRIOR APPLICATION NUMBER: US 08/910,798

PRIOR FILING DATE: 1997-08-13

NUMBER OF SEQ ID NOS: 10

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7

LENGTH: 1901

TYPE: DNA

ORGANISM: Escherichia coli

FEATURE:

NAME/KEY: CDS

LOCATION: (188)...(1483)

FEATURE:

NAME/KEY: misc_feature

LOCATION: 403

OTHER INFORMATION: n = A,T,C or G

US-10-933-115-7

Query Match 97.4%; Score 1287.8; DB 10; Length 1901;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 1293; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 ATGAAAGGATCTTAATCCATTTTATCTCTTGATTTCCGTTAACCCCGCAATCTGCA 60
DB 188 ATGAAAGGATCTTAATCCATTTTATCTCTTGATTTCCGTTAACCCCGCAATCTGCA 247
OY 61 TTGCGTCAGAGTGAGCGGAGCTGAAAGTGAAAGTGATTTGTGATGTCATGATGAT 120
DB 248 TTGCGTCAGAGTGAGCGGAGCTGAAAGTGAAAGTGATTTGTGATGTCATGATGAT 307
OY 121 GTGCGTGTCTCAACCAAGGCCACGCACTGATGTCAGATGTCACCCGACGATGCGCA 180
DB 308 GTGCGTGTCTCAACCAAGGCCACGCACTGATGTCAGATGTCACCCGACGATGCGCA 367
OY 181 ACCTGCGCGGTAAACTGCGTTGCTGACACCGCGAGGTGTGAGCTTAATGCTATCTC 240
DB 368 ACCTGCGCGGTAAACTGCGTTGCTGACACCGCGAGGTGTGAGCTTAATGCTATCTC 427

OY 241 GGAATTACCAAGCGCAGCGCTGTGTAGCCGACGGAATTGCTGCGGAAAAGGCGTCCCG 300
DB 428 GGAATTACCAAGCGCAGCGCTGTGTAGCCGACGGAATTGCTGCGGAAAAGGCGTCCCG 487
OY 301 CAGTCTGCTCAGGTGCGGATTTATGTCGATGTCAGACGCTACCCGTAAACAGCGGAA 360
DB 488 CAGTCTGCTCAGGTGCGGATTTATGTCGATGTCAGACGCTACCCGTAAACAGCGGAA 547
OY 361 GCTTCCGCGCGCGGCGTGGACCTGATCTGTCGAAATTAACCGTACATCCCGGCGATACG 420
DB 548 GCTTCCGCGCGCGGCGTGGACCTGATCTGTCGAAATTAACCGTACATCCCGGCGATACG 607
OY 421 TCCAGTCCCGATCCGTTATTTATCTCTAATAAATGCGCGTTTCCCACTGATACGCG 480
DB 608 TCCAGTCCCGATCCGTTATTTATCTCTAATAAATGCGCGTTTCCCACTGATACGCG 667
OY 481 AACGTACTGAACGATCTCAGACGAGGCGACGAGGTCATTTGCTGACTTTACCGGCGAT 540
DB 668 AACGTACTGAACGATCTCAGACGAGGCGACGAGGTCATTTGCTGACTTTACCGGCGAT 727
OY 541 CGGCAACGCGGTTTCCGGAATGGAACGCGTGTCTTAATTTTCCGCAATCAACTGTGC 600
DB 728 CGGCAACGCGGTTTCCGGAATGGAACGCGTGTCTTAATTTTCCGCAATCAACTGTGC 787
OY 601 CTTAAAGTGAAGAACAGACGAAAGCTGTCATTAACGACGATTAACATCCGAACTC 660
DB 788 CTTAAAGTGAAGAACAGACGAAAGCTGTCATTAACGACGATTAACATCCGAACTC 847
OY 661 AAGGTGAGCGCGCAATGTCATTTAAACGCGTCCGTTAAGCTTCGATCAATCTGACG 720
DB 848 AAGGTGAGCGCGCAATGTCATTTAAACGCGTCCGTTAAGCTTCGATCAATCTGACG 907
OY 721 GAGATTTTCTCCGCAACAAAGCAAGGAAATGCGGAGCGCGGCGGAGGAAAGATCAC 780
DB 908 GAGATTTTCTCCGCAACAAAGCAAGGAAATGCGGAGCGCGGCGGAGGAAAGATCAC 967
OY 781 GATTCAACACAGTGAACACCTTGTCTAAGTTTGATTAACGCGCAATTTTATTTCTCAA 840
DB 968 GATTCAACACAGTGAACACCTTGTCTAAGTTTGATTAACGCGCAATTTTATTTCTCAA 1027
OY 841 CGCAGCGCAAGGTTGCCCGGACCGCGCCACCCCGTTATTTGATTTGATCATGCGACG 900
DB 1028 CGCAGCGCAAGGTTGCCCGGACCGCGCCACCCCGTTATTTGATTTGATCATGCGACG 1087
OY 901 TTGACGCCCATTCACCGGCAAAACAGCGGATGTCGATTAACCACTTCAGTCTG 960
DB 1088 TTGACGCCCATTCACCGGCAAAACAGCGGATGTCGATTAACCACTTCAGTCTG 1147
OY 961 TTATTTGCGGACAGATTAATCTGCAATCTGCGCGGCGCATCTGAGCTCAACTGCG 1020
DB 1148 TTATTTGCGGACAGATTAATCTGCAATCTGCGCGGCGCATCTGAGCTCAACTGCG 1207
OY 1021 ACGCTTCCCGGTCAAGCGGATTAACACCGCGCAGGTGTGTAACCTGTGTTGAAACGCTCG 1080
DB 1208 ACGCTTCCCGGTCAAGCGGATTAACACCGCGCAGGTGTGTAACCTGTGTTGAAACGCTCG 1267
OY 1081 CGTGGCGTAAGCAATTAACAGCAGTGTGATTCAGGTTTCGCTGTCTTCCAGACTTTACAG 1140
DB 1268 CGTGGCGTAAGCAATTAACAGCAGTGTGATTCAGGTTTCGCTGTCTTCCAGACTTTACAG 1327
OY 1141 CAGATGCGTGAATAAAGCGCGCTGTCTTAATTAACGCGCGCGGAGAGGTGAAACTGACC 1200
DB 1328 CAGATGCGTGAATAAAGCGCGCTGTCTTAATTAACGCGCGCGGAGAGGTGAAACTGACC 1387
OY 1201 CTGCGAGGATGTGAAGAGCGAAATGCGAGGGGCAATGTGCTGGAGGTTTAAACGCA 1260
DB 1388 CTGCGAGGATGTGAAGAGCGAAATGCGAGGGGCAATGTGCTGGAGGTTTAAACGCA 1447
OY 1261 ATCGTAATGAAGCAGCATACCGGCGTGCAGTTTGAGAT 1300
DB 1448 ATCGTAATGAAGCAGCATACCGGCGTGCAGTTTGATAT 1487

```
RESULT 11
US-11-056-354-3
; Sequence 3, Application US/11056354
; Publication No. US20050246780A1
; GENERAL INFORMATION:
; APPLICANT: SHORT, Jay M.
; APPLICANT: KREIZ, Keith
; APPLICANT: GRAY, Kevin A.
; APPLICANT: BARTON, Nelson R.
; APPLICANT: GARRETT, James B.
; APPLICANT: O'DONOGHUE, Eileen
; APPLICANT: MAHUR, Eric J.
; TITLE OF INVENTION: PHYASES, NUCLEIC ACIDS ENCODING THEM AND METHODS
; TITLE OF INVENTION: FOR MAKING AND USING THEM
; FILE REFERENCE: 56446201803/D1370-9C1
; CURRENT APPLICATION NUMBER: US/11/056,354
; PRIOR FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 10/156,660
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 3
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli appa phycase
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (403)..(403)
; OTHER INFORMATION: n is any nucleotide
US-11-056-354-3

Query Match          97.4%; Score 1287.8; DB 14; Length 1901;
Beet Local Similarity 99.5%; Pred. No. 0;
Matches 1293; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATGAAGCATCTTAATCCCATTTTATCTTCTGATTCGGTTAACCCCGCAATGCA 60
DB 188 ATGAAGCATCTTAATCCCATTTTATCTTCTGATTCGGTTAACCCCGCAATGCA 247
QY 61 TTCTGCTCAGAGTGAGCCGAGCTGAAAGTGTGATGATGTCAGTGTCTCATGT 120
DB 248 TTCTGCTCAGAGTGAGCCGAGCTGAAAGTGTGATGATGTCAGTGTCTCATGT 307
QY 121 GTGCGTGTCTCAACCAAGGCTCAGCACTGATGACAGATGTCACCCCAAGCGATGCGCA 180
DB 308 GTGCGTGTCTCAACCAAGGCTCAGCACTGATGACAGATGTCACCCCAAGCGATGCGCA 367
QY 181 ACCTGGCCGGTAAACTGGGTTGGCTGACACCGCGGAGTGAGTAAATCGCTATCTC 240
DB 368 ACCTGGCCGGTAAACTGGGTTGGCTGACACCGCGGAGTGAGTAAATCGCTATCTC 427
QY 241 GGAATTTACCAACGCCAGGCTGTGTAGCCGACGATTCGTGGCGAAAAAGGCTGCCG 300
DB 428 GGAATTTACCAACGCCAGGCTGTGTAGCCGACGATTCGTGGCGAAAAAGGCTGCCG 487
QY 301 CAGTCTGTGTCAAGTCCGATTTATGCTGATGTGACAGACGTAACCCGTTAAACAGGCGAA 360
DB 488 CAGTCTGTGTCAAGTCCGATTTATGCTGATGTGACAGACGTAACCCGTTAAACAGGCGAA 547
QY 361 GCCTTGCCGCGGGGTGACCTGACTGTGCAATTAACGTATACCAAGCAAGATAG 420
```

```
DB 548 GCCTTGCCGCGGGGTGACCTGACTGTGCAATTAACGTATACCAAGCAAGATAG 607
QY 421 TCAGTCCCGATCCGTTATTTATCTCTTAAAACTGGCGTTTCCCACTGATTAACGG 480
DB 608 TCAGTCCCGATCCGTTATTTATCTCTTAAAACTGGCGTTTCCCACTGATTAACGG 667
QY 481 AACGTACTGACGCGATCTCTGACAGGAGGAGGATGAAATGTCGACTTTACCGGGCAT 540
DB 668 AACGTACTGACGCGATCTCTGACAGGAGGAGGATGAAATGTCGACTTTACCGGGCAT 727
QY 541 CGGCAACCGCGTTTTCGCAACTGNAACGGGTCTTAATTTTCCGCAATCAACTGTGTC 600
DB 728 CGGCAACCGCGTTTTCGCAACTGNAACGGGTCTTAATTTTCCGCAATCAACTGTGTC 787
QY 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTACGACGACATTCATTCGAACTC 660
DB 788 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTACGACGACATTCATTCGAACTC 847
QY 661 AAGGTGAGCGCGACAAATGTCATTAAACGGTGGGTAAAGCTCGCAATCAATGTGAG 720
DB 848 AAGGTGAGCGCGACAAATGTCATTAAACGGTGGGTAAAGCTCGCAATCAATGTGAG 907
QY 721 GAGATATTTCTCTGCAACAAGCAAGGATATGCGAGCCGGGTGGGAAAGATCAC 780
DB 908 GAGATATTTCTCTGCAACAAGCAAGGATATGCGAGCCGGGTGGGAAAGATCAC 967
QY 781 GATTCACACCAAGTGAACACCTTGCTAAGTTTGATTAACGCGCAATTTTATTTGCTACA 840
DB 968 GATTCACACCAAGTGAACACCTTGCTAAGTTTGATTAACGCGCAATTTTATTTGCTACA 1027
QY 841 CGCAGCGCAAGAGTTGCGCGAGCCGCGCACCCCGTTATGATTTGATTCATGAGAG 900
DB 1028 CGCAGCGCAAGAGTTGCGCGAGCCGCGCACCCCGTTATGATTTGATTCATGAGAG 1087
QY 901 TTGAGCGCCCATCAACCGCAAAAAAGCGGTATGATGATCAATCCCACTTCAAGTCTG 960
DB 1088 TTGAGCGCCCATCAACCGCAAAAAAGCGGTATGATGATCAATCCCACTTCAAGTCTG 1147
QY 961 TTTATTTCCGGAACAGATATCTAATTTGGCAATCTCGCGCGGCACTGAGGCTCAACTGG 1020
DB 1148 TTTATTTCCGGAACAGATATCTAATTTGGCAATCTCGCGCGGCACTGAGGCTCAACTGG 1207
QY 1021 AAGCTTCCCGGTCAACCGGATTAACAAGCCGCAAGGTGATGAACTGTGTTAAAGCTGG 1080
DB 1208 AAGCTTCCCGGTCAACCGGATTAACAAGCCGCAAGGTGATGAACTGTGTTAAAGCTGG 1267
QY 1081 CGTGGCTTAAGGATTAACGCAAGTGAATTCAGTTTGGCTGCTTCCAGACTTTACG 1140
DB 1268 CGTGGCTTAAGGATTAACGCAAGTGAATTCAGTTTGGCTGCTTCCAGACTTTACG 1327
QY 1141 CAGATGCGTGAATTAACGCGCTGTCTAATTAATACGCGCCCGGAGAGTGAATGAC 1200
DB 1328 CAGATGCGTGAATTAACGCGCTGTCTAATTAATACGCGCCCGGAGAGTGAATGAC 1387
QY 1201 CTGGCAGATGTAAGAGCGAAATGCGCAGGCGCATGTGTTGTTGCAAGTTTACGCA 1260
DB 1388 CTGGCAGATGTAAGAGCGAAATGCGCAGGCGCATGTGTTGTTGCAAGTTTACGCA 1447
QY 1261 ATCGTAATGAAGACGCAATACCGGCTGCAAGTTGAAGT 1300
DB 1448 ATCGTAATGAAGACGCAATACCGGCTGCAAGTTGAAGT 1487

RESULT 12
US-10-282-122A-7167
; Sequence 7167, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
```

APPLICANT: Zykkind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: EITRA.034A
CURRENT FILING DATE: US/10/282,122A
PRIOR APPLICATION NUMBER: 2003-02-20
PRIOR FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7167
LENGTH: 1299
TYPE: DNA
ORGANISM: Escherichia coli
US-10-282-122A-7167

Query Match 97.4%; Score 1287; DB 8; Length 1299;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1290; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

1 ATGAAGGAGTCTTATCCATTTTATCTCTCGATTCGTTAAACCCCGAATCTCA 60
1 ATGAAGGAGTCTTATCCATTTTATCTCTCGATTCGTTAAACCCCGAATCTCA 60
61 TTCGCTCAGAGTGAACCGGAGCTGAAGCTGAAAGTGTGTAATTGTCAGTGTGATGAT 120
61 TTCGCTCAGAGTGAACCGGAGCTGAAGCTGAAAGTGTGTAATTGTCAGTGTGATGAT 120
121 GTGCGTGTCCAAACCAAGGCCACGCACTGATGAGATGTCAACCCGAGCGATGAGCA 180
121 GTGCGTGTCCAAACCAAGGCCACGCACTGATGAGATGTCAACCCGAGCGATGAGCA 180
121 GTGCGTGTCCAAACCAAGGCCACGCACTGATGAGATGTCAACCCGAGCGATGAGCA 180
121 GTGCGTGTCCAAACCAAGGCCACGCACTGATGAGATGTCAACCCGAGCGATGAGCA 180
181 ACCTGGCCGGTAAACTGTGGTGTGTCACACCGCGAGTGTGTAAGTGTGTAAGTGTG 240
181 ACCTGGCCGGTAAACTGTGGTGTGTCACACCGCGAGTGTGTAAGTGTGTAAGTGTG 240
181 ACCTGGCCGGTAAACTGTGGTGTGTCACACCGCGAGTGTGTAAGTGTGTAAGTGTG 240
241 GGAACATTAACAAGCGACGCTGTGTGTAAGCGAGATTGCTGGCGAAGGAGGCTGCCG 300
241 GGAACATTAACAAGCGACGCTGTGTGTAAGCGAGATTGCTGGCGAAGGAGGCTGCCG 300
301 CAGTGTGTCAAGTGTGCGATTATTTGCTGATGTGCAAGCGTAACCCGTAACCAAGCGAA 360
301 CAGTGTGTCAAGTGTGCGATTATTTGCTGATGTGCAAGCGTAACCCGTAACCAAGCGAA 360
361 GCTTTCGCGCGCGGCTGGACCTGATCTGTGAATTAACCGTAACCAAGCGAATATACG 420
361 GCTTTCGCGCGCGGCTGGACCTGATCTGTGAATTAACCGTAACCAAGCGAATATACG 420
421 TCAGTCCGAGTCCGTTATTTATCTCTCTAAACCTGGGCTTTGCAACTGAGTAACCGG 480

|||||
421 TCAGTCCGAGTCCGTTATTTATCTCTCTAAACCTGGGCTTTGCAACTGAGTAACCGG 480
481 AACGTACTGACGCGATCTCTAGACAGGCGAGAGGTCATTTGCTGACTTTACCGGGCAT 540
481 AACGTACTGACGCGATCTCTAGACAGGCGAGAGGTCATTTGCTGACTTTACCGGGCAT 540
541 CCGCAACCGGCTTTCGGAACCTGAAACGGGTGCTTAATTTTCGCAATCAAACTTGTC 600
541 CCGCAACCGGCTTTCGGAACCTGAAACGGGTGCTTAATTTTCGCAATCAAACTTGTC 600
601 CTTAAACGTGAAACAGGAGGAGCTGTTCAATTAACGAGGATTAACCATCGAACTC 660
601 CTTAAACGTGAAACAGGAGGAGCTGTTCAATTAACGAGGATTAACCATCGAACTC 660
661 AAGGTAGCGCGGACATGTCTCAATTAACCGGTGCGGTGAGCTTCGATCAATGCTGACG 720
661 AAGGTAGCGCGGACATGTCTCAATTAACCGGTGCGGTGAGCTTCGATCAATGCTGACG 720
721 GAGATATTTCTCTGCAACAGCAAGGAAATGCGGAGCGGGGTGGGAAAGATCAC 780
721 GAGATATTTCTCTGCAACAGCAAGGAAATGCGGAGCGGGGTGGGAAAGATCAC 780
781 GATTCAACCAAGTGAACACCTTCTAAATTGCAATACGCGCAATTTATTTGCTACAA 840
781 GATTCAACCAAGTGAACACCTTCTAAATTGCAATACGCGCAATTTATTTGCTACAA 840
841 CGCAGCGGAGGTTGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
841 CGCAGCGGAGGTTGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
901 TTGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
901 TTGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
961 TTTATTCGCGGACAGTACTAATCTGCAATCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
961 TTTATTCGCGGACAGTACTAATCTGCAATCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
1021 ACGCTTCCCGGTCAGCGCGGATTAACAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
1021 ACGCTTCCCGGTCAGCGCGGATTAACAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
1081 CGTCGCGTAAAGCATTAACAGCGGATTAACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
1081 CGTCGCGTAAAGCATTAACAGCGGATTAACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
1141 CAGATTCGCTGATTAACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
1141 CAGATTCGCTGATTAACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
1201 CTGGCAGGATGTGAAGCGGAAATGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1260
1201 CTGGCAGGATGTGAAGCGGAAATGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1260
1261 ATCGTAATGAAGCAGCATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1296
1261 ATCGTAATGAAGCAGCATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1296

RESULT 13
US-09-866-379-5
Sequence 5, Application US/09866379
Patent No. US20020136754A1
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: SHORT, Jay
APPLICANT: KRETZ, Keith
APPLICANT: GRAY, Kevin
APPLICANT: BARTON, Nelson
APPLICANT: GARRETT, James
APPLICANT: O'DONOGHUE, Eileen
TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF

```
/ FILE REFERENCE: DIVER1370-7
/ CURRENT APPLICATION NUMBER: US/09/866,379
/ CURRENT FILING DATE: 2001-05-24
/ PRIOR APPLICATION NUMBER: US 09/580,515
/ PRIOR FILING DATE: 2000-05-25
/ PRIOR APPLICATION NUMBER: US 09/318,528
/ PRIOR FILING DATE: 1999-05-25
/ PRIOR APPLICATION NUMBER: US 09/291,931
/ PRIOR FILING DATE: 1999-04-13
/ PRIOR APPLICATION NUMBER: US 09/259,214
/ PRIOR FILING DATE: 1999-03-01
/ PRIOR APPLICATION NUMBER: US 08/910,798
/ PRIOR FILING DATE: 1997-08-13
/ NUMBER OF SEQ ID NOS: 10
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 5
/ LENGTH: 1901
/ TYPE: DNA
/ ORGANISM: Escherichia coli
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)..(1901)
/ OTHER INFORMATION: n is any nucleotide
US-09-866-379-5
```

Query Match 97.3%; Score 1286.2; DB 3; Length 1901;

Best Local Similarity 99.4%; Pred. No. 0; Matches 1292; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
QY 1 ATGAAAGCATCTTAATCCATTTTATCTCTGATTCCTGTAACCCCGCAATCTGCA 60
DB 188 ATGAAAGCATCTTAATCCATTTTATCTCTGATTCCTGTAACCCCGCAATCTGCA 247
QY 61 TTGCTCAGATGAGCGCGAGCTGAAAGTGTGTGTAATGTCACTGCTCATGTG 120
DB 248 TTGCTCAGATGAGCGCGAGCTGAAAGTGTGTGTAATGTCACTGCTCATGTG 307
QY 121 GTGCGGTCTCCCAACCAAGGCCACGCACTGATGCAAGATGTCAACCCCAAGCGATGGCA 180
DB 308 GTGCGGTCTCCCAACCAAGGCCACGCACTGATGCAAGATGTCAACCCCAAGCGATGGCA 367
QY 181 ACCCTGCGCGGTAAACTGGGTTGGCTGACACCGCGNGTGTGTGTAATGTCACTATCTC 240
DB 368 ACCCTGCGCGGTAAACTGGGTTGGCTGACACCGCGNGTGTGTGTAATGTCACTATCTC 427
QY 241 GGAATTAACCAACGCGAGCGTGTGTAAGCCGCAAGATTGCTGCGCAAAAAAGGCTGCGCG 300
DB 428 GGAATTAACCAACGCGAGCGTGTGTAAGCCGCAAGATTGCTGCGCAAAAAAGGCTGCGCG 487
QY 301 CAGTCTGTGATGAGTGGCGATTTATGTGTATGTGACAGACGTAACCGTAAACAGGCGAA 360
DB 488 CAGTCTGTGATGAGTGGCGATTTATGTGTATGTGACAGACGTAACCGTAAACAGGCGAA 547
QY 361 GCCTTCGCGCGCGGCTGACCTGACCTGCAATTAACGTAATCCAGGCAAGATAG 420
DB 548 GCCTTCGCGCGCGGCTGACCTGACCTGCAATTAACGTAATCCAGGCAAGATAG 607
QY 421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAACTGGGCTTGGCAATGATTAACGCG 480
DB 608 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAACTGGGCTTGGCAATGATTAACGCG 667
QY 481 AACGTAAGTGAAGCGATCTCTGACAGAGGAGGAGGATCAATTGCTGACTTTAACCGGCGAT 540
DB 668 AACGTAAGTGAAGCGATCTCTGACAGAGGAGGAGGATCAATTGCTGACTTTAACCGGCGAT 727
QY 541 CGGCAAAACGGCTTTGCGCAACTGAAACGGGTGCTTAATTTTCGCAATCAAACTTGTGC 600
DB 728 CGGCAAAACGGCTTTGCGCAACTGAAACGGGTGCTTAATTTTCGCAATCAAACTTGTGC 787
QY 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAAACGACAGCAATTACATCGAATCTC 660
DB 788 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAAACGACAGCAATTACATCGAATCTC 847
```

```
QY 661 AAGTGAAGCGCGCAACATCTTAAACCGGTGCGGTAAAGCTCGCATCAATGCTGAGC 720
DB 848 AAGTGAAGCGCGCAACATCTTAAACCGGTGCGGTAAAGCTCGCATCAATGCTGAGC 907
QY 721 GAGATATTTCTCCTGCAACAGACAGGAAATGCGGACCGGGTGGGAAAGATCAAC 780
DB 908 GAGATATTTCTCCTGCAACAGACAGGAAATGCGGACCGGGTGGGAAAGATCAAC 967
QY 781 GATTCAACCAAGTGAACCACTTGTCTAAGTTTGATTAACCGCAATTTTATTTGCTAAC 840
DB 968 GATTCAACCAAGTGAACCACTTGTCTAAGTTTGATTAACCGCAATTTTATTTGCTAAC 1027
QY 841 CGACGCGCAGAGTGGCCCGGACCGCGGCAACCCGTTATTTGATTTGATGAGGAGC 900
DB 1028 CGACGCGCAGAGTGGCCCGGACCGCGGCAACCCGTTATTTGATTTGATGAGGAGC 1087
QY 901 TTGAACCCCATCAACCGCAAAACAGGCGTATGATGATCAATTAACCACTTCACTAG 960
DB 1088 TTGAACCCCATCAACCGCAAAACAGGCGTATGATGATCAATTAACCACTTCACTAG 1147
QY 961 TTTATTTGCGGACACGATATCTAATCTGCAAACTTGGCGGCGCACTGAGCTCAACTG 1020
DB 1148 TTTATTTGCGGACACGATATCTAATCTGCAAACTTGGCGGCGCACTGAGCTCAACTG 1207
QY 1021 AAGCTTCCCGGTCAAGCGGATTAACAGCCGCGCAAGTGTGAACTGTGTTTGAACGCTG 1080
DB 1208 AAGCTTCCCGGTCAAGCGGATTAACAGCCGCGCAAGTGTGAACTGTGTTTGAACGCTG 1267
QY 1081 CGTGGCTTAAGGATTAACAGCGAGTGAATTCAGTTTGTGCTGCTTCAAGCTTAAACG 1140
DB 1268 CGTGGCTTAAGGATTAACAGCGAGTGAATTCAGTTTGTGCTGCTTCAAGCTTAAACG 1327
QY 1141 CAGATGCGTGAATAAACCGCGTGTCTATTAAATACCGCGCCGAGAGGTGAACCTGAC 1200
DB 1328 CAGATGCGTGAATAAACCGCGTGTCTATTAAATACCGCGCCGAGAGGTGAACCTGAC 1387
QY 1201 CTGGCAGATGTGAAGAGGAAATGCGCAGGGCATGTGTTGTTGGCAGGTTTACGAA 1260
DB 1388 CTGGCAGATGTGAAGAGGAAATGCGCAGGGCATGTGTTGTTGGCAGGTTTACGAA 1447
QY 1261 ATCGTAATGAAGACGCACTACCGGCGTGCAGTTTGAGAT 1300
DB 1448 ATCGTAATGAAGACGCACTACCGGCGTGCAGTTTGAGAT 1487

RESULT 14
US-09-866-379-9
/ Sequence 9, Application US/09866379
/ Patent No. US20020136754A1
/ GENERAL INFORMATION:
/ APPLICANT: DIVERSA CORPORATION
/ APPLICANT: SHORT, Jay
/ APPLICANT: KRETTZ, Keith
/ APPLICANT: GRAY, Kevin
/ APPLICANT: BARTON, Nelson
/ APPLICANT: GARRETT, James
/ APPLICANT: O'DONOGHUE, Rileen
/ TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
/ FILE REFERENCE: DIVER1370-7
/ CURRENT APPLICATION NUMBER: US/09/866,379
/ CURRENT FILING DATE: 2001-05-24
/ PRIOR APPLICATION NUMBER: US 09/580,515
/ PRIOR FILING DATE: 2000-05-25
/ PRIOR APPLICATION NUMBER: US 09/318,528
/ PRIOR FILING DATE: 1999-05-25
/ PRIOR APPLICATION NUMBER: US 09/291,931
/ PRIOR FILING DATE: 1999-04-13
/ PRIOR APPLICATION NUMBER: US 09/259,214
/ PRIOR FILING DATE: 1999-03-01
/ PRIOR APPLICATION NUMBER: US 08/910,798
/ PRIOR FILING DATE: 1997-08-13
/ NUMBER OF SEQ ID NOS: 10
/ SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 9
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-9
```

```
Query Match      97.3%; Score 1286.2; DB 3; Length 1901;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1292; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
QY      1 ATGAAAGGATCTTATCCATTTTATCTCTTGATTCGTTACCCCGCAATCTGA 60
DB      188 ATGAAAGGATTTTATCCATTTTATCTCTTGATTCGTTACCCCGCAATCTGA 247
QY      61 TTCGCTCAGATGAGCCGAGCTGAAAGCTGAAAGTGTGATTTGTCACTGTCATGAT 120
DB      248 TTCGCTCAGATGAGCCGAGCTGAAAGTGTGATTTGTCACTGTCATGATGAT 307
QY      121 GTGCGTCTCCACCAAGCCCAAGCACTGATGACGATGTCACTCCAGAGCATGGCCA 180
DB      308 GTGCGTCTCCACCAAGCCCAAGCACTGATGACGATGTCACTCCAGAGCATGGCCA 367
QY      181 ACCTGGCGGTAAACTGGGTTGGCTGACACCGCGNGGTGGTGAAGCTATGCGCTATCTC 240
DB      368 ACTGGCGGTAAACTGGGTTGGCTGACACCGCGNGGTGGTGAAGCTATGCGCTATCTC 427
QY      241 GGACATTAACAACCGCAGCGTCTGTTAGCCGACGATTTGCTGGCCGAAAGGGCTGCCCG 300
DB      428 GGACATTAACAACCGCAGCGTCTGTTAGCCGACGATTTGCTGGCCGAAAGGGCTGCCCG 487
QY      301 CAGTCTGTCAGTGTGCGCATTTTCTGATGTTCGACGAGCGTATCCCGTAAACAGGCGAA 360
DB      488 CAGTCTGTCAGTGTGCGCATTTTCTGATGTTCGACGAGCGTATCCCGTAAACAGGCGAA 547
QY      361 GCCTTCGCGCGCGGGCTGGCACTGACGTGCAATTAACCGTACATACCAGGCAATACG 420
DB      548 GCCTTCGCGCGCGGGCTGGCACTGACGTGCAATTAACCGTACATACCAGGCAATACG 607
QY      421 TCCAGTCCCGATCCGTTATTTATCTCTTAATAAACTGGCGTTTGGCAACTGATTAACCG 480
DB      608 TCCAGTCCCGATCCGTTATTTATCTCTTAATAAACTGGCGTTTGGCAACTGATTAACCG 667
QY      481 AACGTGATGACGCGATCTCTGACGAGGCGCAAGAGGCTCAATTGCTGACTTTAACCGGGCAT 540
DB      668 AACGTGATGACGCGATCTCTGACGAGGCGCAAGAGGCTCAATTGCTGACTTTAACCGGGCAT 727
QY      541 CGGCAAAAGGGGTTTCGGAACTGGAACGGGTCTTAATTTTCGGCATCAAACTTGTGC 600
DB      728 CGGCAAAAGGGGTTTCGGAACTGGAACGGGTCTTAATTTTCGGCATCAAACTTGTGC 787
QY      601 CTTAAACGTGAAAAACAGACGAAAGCTGTTCAATTAAACGACGACATTAACATCGGAATC 660
DB      788 CTTAAACGTGAAAAACAGACGAAAGCTGTTCAATTAAACGACGACATTAACATCGGAATC 847
QY      661 AAGGTGAGCGCGCAATGTCTCATTAACCGGTGGGTGAAGCTTGCATCAATGCTGACG 720
DB      848 AAGGTGAGCGCGCAATGTCTCATTAACCGGTGGGTGAAGCTTGCATCAATGCTGACG 907
QY      721 GAGATATTTCTCTGCAACAAGCAAGGGAATGCGGAGCGGGGTGGGAAAGATCAC 780
DB      908 GAGATATTTCTCTGCAACAAGCAAGGGAATGCGGAGCGGGGTGGGAAAGATCAC 967
QY      781 GATTTCACACGATGGAACCTTGTCTAAGTTGCAATTAACGCGCAATTTTATTTGCTACAA 840
DB      968 GATTTCACACGATGGAACCTTGTCTAAGTTGCAATTAACGCGCAATTTTATTTGCTACAA 1027
QY      841 CGCAGCGCAAGAGTTGCCCGCAGCGCGCCACCCCGTTATTTGATTTGATTCATGCGACGG 900
DB      1028 CGCAGCGCAAGAGTTGCCCGCAGCGCGCCACCCCGTTATTTGATTTGATTCATGCGACGG 1087
```

```
QY      901 TTGACGCCCCATCCACCGCAAAAAACAGCGGATGATGTGACATTACCATTCACTTCACTG 960
DB      1088 TTGACGCCCCATCCACCGCAAAAAACAGCGGATGATGTGACATTACCATTCACTTCACTG 1147
QY      961 TTTATTTGCGCGGACACGATTAATCTGTGGCAAAATCTCGCGGGCGCACTGAGCTCACTGG 1020
DB      1148 TTTATTTGCGCGGACACGATTAATCTGTGGCAAAATCTCGCGGGCGCACTGAGCTCACTGG 1207
QY      1021 ACGCTTCCCGGTCAAGCGGATTAACAAGCCGCGCAGGTGTGAACGTGTGTAAGCGCTGG 1080
DB      1208 ACGCTTCCCGGTCAAGCGGATTAACAAGCCGCGCAGGTGTGAACGTGTGTAAGCGCTGG 1267
QY      1081 CGTCGCTAAGCAGATTAACAGCCAGTGAATTCAGGTTTCGCTGTCTTCCAGACTTTACAG 1140
DB      1268 CGTCGCTAAGCAGATTAACAGCCAGTGAATTCAGGTTTCGCTGTCTTCCAGACTTTACAG 1327
QY      1141 CAGATGCGTGAATAAACCGCGCTGTCAATTAAATACCGCGCCGAGAGGTGAACCTGACC 1200
DB      1328 CAGATGCGTGAATAAACCGCGCTGTCAATTAAATACCGCGCCGAGAGGTGAACCTGACC 1387
QY      1201 CTGGCAGGATGTGAAGACGGAATATGCCAGAGGCAATGTTGTTGGCAGGTTTAAACGCA 1260
DB      1388 CTGGCAGGATGTGAAGACGGAATATGCCAGAGGCAATGTTGTTGGCAGGTTTAAACGCA 1447
QY      1261 ATCGTAATGAAGCAGCATACCGGCGTGCAGTTTGAGAT 1300
DB      1448 ATCGTAATGAAGCAGCATACCGGCGTGCAGTTTGAT 1487
```

```
RESULT 15
US-10-601-319-5
```

```
; Sequence 5, Application US/10601319
; Publication No. US20040091968A1
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Short, Jay M.
```

```
; APPLICANT: Kretz, Keith A.
```

```
; APPLICANT: Gray, Kevin A.
```

```
; APPLICANT: Barton, Nelson Robert
```

```
; APPLICANT: Garrett, James B.
```

```
; APPLICANT: O'Donoghue, Bileen
```

```
; TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
```

```
; FILE REFERENCE: 09010-029011
```

```
; CURRENT APPLICATION NUMBER: US/10/601,319
```

```
; CURRENT FILING DATE: 2003-06-20
```

```
; PRIOR APPLICATION NUMBER: US 09/866,379
```

```
; PRIOR FILING DATE: 2001-05-24
```

```
; PRIOR APPLICATION NUMBER: US 09/580,515
```

```
; PRIOR FILING DATE: 2000-05-25
```

```
; PRIOR APPLICATION NUMBER: US 09/318,528
```

```
; PRIOR FILING DATE: 1999-05-25
```

```
; PRIOR APPLICATION NUMBER: US 09/291,931
```

```
; PRIOR FILING DATE: 1999-04-13
```

```
; PRIOR APPLICATION NUMBER: US 09/259,214
```

```
; PRIOR FILING DATE: 1999-03-01
```

```
; PRIOR APPLICATION NUMBER: US 08/910,798
```

```
; PRIOR FILING DATE: 1997-08-13
```

```
; NUMBER OF SEQ ID NOS: 10
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 5
```

```
; LENGTH: 1901
```

```
; TYPE: DNA
```

```
; ORGANISM: Escherichia coli
```

```
; FEATURE:
```

```
; NAME/KEY: misc_feature
```

```
; LOCATION: 403
```

```
; OTHER INFORMATION: n = A,T,C or G
```

```
US-10-601-319-5
```

```
Query Match      97.2%; Score 1284.6; DB 8; Length 1901;
Best Local Similarity 99.3%; Pred. No. 0;
```

Matches 1291; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```
QY 1 ATGAAGGCACTTTATCCCATTTTATCTCTTCTGATTTCCGTTAACCCCGCATATGCA 60
Db 188 ATGAAGGCACTTTATCCCATTTTATCTCTTCTGATTTCCGTTAACCCCGCATATGCA 247
QY 61 TTGCGCTCAGAGTGAAGCCGAGCTGAAGCTGAAGAGTGTGTGATTTGTCAGTCCGTAAGT 120
Db 248 TTGCGCTCAGAGTGAAGCCGAGCTGAAGAGTGTGTGATTTGTCAGTCCGTAAGT 307
QY 121 GTGCGTCTCCACCAAGGCGACGCAATGTATGACAGATGTCACCCGAGACGATGCGCA 180
Db 308 GTGCGTCTCCACCAAGGCGACGCAATGTATGACAGATGTCACCCGAGACGATGCGCA 367
QY 181 ACCTGCCCGGTAAACTGTGGTGTGTGATCAACCGGAGGTGTGTGAGCTAATCGCTATCTC 240
Db 368 ACCTGCCCGGTAAACTGTGGTGTGTGATCAACCGGAGGTGTGTGAGCTAATCGCTATCTC 427
QY 241 GGCATTATCAACGCGAGCGTGTGAGCGACGCAATGTCGCGAAGAGGCTGCGCG 300
Db 428 GGCATTATCAACGCGAGCGTGTGAGCGACGCAATGTCGCGAAGAGGCTGCGCG 487
QY 301 CAGTCTGTCAAGTGTGCGATTTATGTCTGATGTGACAGAGGTACCCGTAAACAGGCGAA 360
Db 488 CAGTCTGTCAAGTGTGCGATTTATGTCTGATGTGACAGAGGTACCCGTAAACAGGCGAA 547
QY 361 GCTTGTGCGCGCGGCTGTGCACTGTGCAATACCGTACATACCGAGCATACG 420
Db 548 GCTTGTGCGCGCGGCTGTGCACTGTGCAATACCGTACATACCGAGCATACG 607
QY 421 TCCAGTCCCGATCCGTTATTTAACTCTTAAACCTGGGTTGGCACTGTATACGCG 480
Db 608 TCCAGTCCCGATCCGTTATTTAACTCTTAAACCTGGGTTGGCACTGTATACGCG 667
QY 481 AACGTGACTGACGCGATCTCTCAAGCAGGAGGAGGCTCAATTGCTGACTTTACCGGCGAT 540
Db 668 AACGTGACTGACGCGATCTCTCAAGCAGGAGGAGGCTCAATTGCTGACTTTACCGGCGAT 727
QY 541 CCGCAAAACGCGGCTTTGCGCACTGTGAACGGGTGCTTAATTTTCCGCAATCAACTTGTGC 600
Db 728 CCGCAAAACGCGGCTTTGCGCACTGTGAACGGGTGCTTAATTTTCCGCAATCAACTTGTGC 787
QY 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTCAATTAAAGCAGGCAATTACATCGAACTC 660
Db 788 CTTAAACGTGAGAAACAGGACGAAAGCTGTCAATTAAAGCAGGCAATTACATCGAACTC 847
QY 661 AAGGTGAGCGCGCAATGTCTCATTTAACGGGTGCGTAAAGCTCGCATCAATGCTGACG 720
Db 848 AAGGTGAGCGCGCAATGTCTCATTTAACGGGTGCGTAAAGCTCGCATCAATGCTGACG 907
QY 721 GAGATATTTCTCTGTCAACAGACAGGGAATGCCGAGCCGCGGTGGGAAAGATCAC 780
Db 908 GAGATATTTCTCTGTCAACAGACAGGGAATGCCGAGCCGCGGTGGGAAAGATCAC 967
QY 781 GATTCACACAGTGAACACCTTGTCAATGTTTGCATTAAGCGCAATTTATTTGCTACAA 840
Db 968 GATTCACACAGTGAACACCTTGTCAATGTTTGCATTAAGCGCAATTTATTTGCTACAA 1027
QY 841 CGCAGCGCAGAGGTTGCCGCGACGCGCGCAACCCGTTATTGGAATTTGATGAGCAGCG 900
Db 1028 CGCAGCGCAGAGGTTGCCGCGACGCGCGCAACCCGTTATTGGAATTTGATGAGCAGCG 1087
QY 901 TTGAGCGCCCATTCACCGCAAAACAGGCGTATGTGTGACATTAACCACTTCAAGTCTG 960
Db 1088 TTGAGCGCCCATTCACCGCAAAACAGGCGTATGTGTGACATTAACCACTTCAAGTCTG 1147
QY 961 TTTATTTGCGGACAGATTAATCTGGAATATCTCGGGGCGCACTGAGAGTCAACTGG 1020
Db 1148 TTTATTTGCGGACAGATTAATCTGGAATATCTCGGGGCGCACTGAGAGTCAACTGG 1207
QY 1021 ACGCTTCCCGGTCAAGCCGATTAACGCGCGAGGTGTGAATCTGGTGTGAAACGCTGG 1080
Db 1208 ACGCTTCCCGGTCAAGCCGATTAACGCGCGAGGTGTGAATCTGGTGTGAAACGCTGG 1267
```

```
QY 1081 CGTGCGCTAAGCGATTAACGCCAGTGTGATTCAGGTTTGTGCTTCCAGACTTTACAG 1140
Db 1268 CGTGCGCTAAGCGATTAACGCCAGTGTGATTCAGGTTTGTGCTTCCAGACTTTACAG 1327
QY 1141 CAGATGCGTATTAACGCGCGTGTCAATTAAATACGCGCCGCGAGAGGTGAACAGTAC 1200
Db 1328 CAGATGCGTATTAACGCGCGTGTCAATTAAATACGCGCCGCGAGAGGTGAACAGTAC 1387
QY 1201 CTGGCAGATGTGAAGAGCGAAATGCGCAGGCGATGTGTGTTGTCAGAGTTTACGCA 1260
Db 1388 CTGGCAGATGTGAAGAGCGAAATGCGCAGGCGATGTGTGTTGTCAGAGTTTACGCA 1447
QY 1261 ATGTGAAATGAAGCAGCATACCGGCTGCAAGTTTGAAT 1300
Db 1448 ATGTGAAATGAAGCAGCATACCGGCTGCAAGTTTGAAT 1487
```

Search completed: June 14, 2006, 12:41:35
Job time : 1618.16 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2006, 10:04:50 ; Search time 263.494 Seconds
(without alignments)
9394.823 Million cell updates/sec

Title: US-10-601-319-1

Perfect score: 1322
Sequence: 1 atgaagagcattcaatcc.....atcaccatcaccatcaaa 1323

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/ina/1_COMB.seq:*
- 2: /EMC_Celerra_SIDS3/ptodata/2/ina/5_COMB.seq:*
- 3: /EMC_Celerra_SIDS3/ptodata/2/ina/6A_COMB.seq:*
- 4: /EMC_Celerra_SIDS3/ptodata/2/ina/6B_COMB.seq:*
- 5: /EMC_Celerra_SIDS3/ptodata/2/ina/7_COMB.seq:*
- 6: /EMC_Celerra_SIDS3/ptodata/2/ina/H_COMB.seq:*
- 7: /EMC_Celerra_SIDS3/ptodata/2/ina/PCpus_COMB.seq:*
- 8: /EMC_Celerra_SIDS3/ptodata/2/ina/PP_COMB.seq:*
- 9: /EMC_Celerra_SIDS3/ptodata/2/ina/RE_COMB.seq:*
- 10: /EMC_Celerra_SIDS3/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1322	100.0	1323	US-09-259-214-1	Sequence 1, Appl1
2	1322	100.0	1323	US-09-318-528-1	Sequence 1, Appl1
3	1322	100.0	1323	US-09-291-931-1	Sequence 1, Appl1
4	1322	100.0	1323	US-09-580-515-1	Sequence 1, Appl1
5	1322	100.0	1323	US-09-866-379-1	Sequence 1, Appl1
6	1287.8	97.4	1301	US-09-866-379-7	Sequence 5, Appl1
7	1286.2	97.3	1301	US-09-866-379-5	Sequence 5, Appl1
8	1283	97.0	1301	US-09-866-379-9	Sequence 6, Appl1
9	1283	95.6	1489	US-09-540-149A-9	Sequence 9, Appl1
10	1263.8	95.6	1489	US-10-266-041A-9	Sequence 9, Appl1
11	1263.8	95.5	1486	US-09-715-477-2	Sequence 2, Appl1
12	1263	94.7	1486	US-09-715-477-4	Sequence 4, Appl1
13	1251.8	94.7	1486	US-08-910-798-1	Sequence 4, Appl1
14	1210	91.5	1272	US-08-910-798-1	Sequence 4, Appl1
15	63.6	4.8	1266	US-09-489-039A-341	Sequence 341, App
16	39.8	3.0	2868	US-09-710-794-4	Sequence 15917, A
17	36.4	2.8	480	US-09-252-991A-15917	Sequence 15888, A
18	36.4	2.8	762	US-09-252-991A-15888	Sequence 15774, A
19	36.4	2.8	885	US-09-252-991A-15774	Sequence 15804, A
20	34.6	2.6	1068	US-09-252-991A-15804	Sequence 15102, Ap
21	34.6	2.6	14570	US-09-902-540-1012	Sequence 1012, Ap
22	34.6	2.6	3076	US-09-710-794-1	Sequence 1, Appl1
23	34.4	2.6	3076	US-09-710-794-1	Sequence 1, Appl1

24	34.4	2.6	4403765	3	US-09-103-840A-2	Sequence 2, Appl1
25	34.4	2.6	4411529	3	US-09-103-840A-1	Sequence 5, Appl1
26	34	2.6	998	2	US-07-885-089B-5	Sequence 8907, Ap
27	34	2.6	1044	3	US-09-252-991A-8907	Sequence 8696, Ap
28	34	2.6	1068	3	US-09-710-794-3	Sequence 9082, Ap
29	34	2.6	2113	3	US-09-252-991A-6696	Sequence 16384, A
30	34	2.6	2706	3	US-09-252-991A-9082	Sequence 583, App
31	33.8	2.6	12257	3	US-09-949-016-16384	Sequence 8976, Ap
32	33.6	2.5	7515	3	US-09-949-016-583	Sequence 15639, A
33	33.4	2.5	399	3	US-09-621-976-8976	Sequence 14877, App
34	33.2	2.5	505	3	US-09-621-976-15639	Sequence 116, App
35	33.2	2.5	1584	3	US-09-489-039A-497	Sequence 17, Appl
36	33	2.5	505	3	US-09-621-976-15639	Sequence 17, Appl
37	33	2.5	2511	2	US-08-680-328-116	Sequence 14187, A
38	33	2.5	35100	2	US-08-770-379-17	Sequence 16326, A
39	33	2.5	35100	3	US-08-757-669A-17	Sequence 17402, A
40	33	2.5	35100	3	US-09-230-371A-17	Sequence 2259, Ap
41	33	2.5	69701	3	US-09-949-016-14187	Sequence 2538, Ap
42	33	2.5	73308	3	US-09-949-016-16326	
43	33	2.5	84425	3	US-09-949-016-17402	
44	32.8	2.5	1362	3	US-09-489-039A-2259	
45	32.8	2.5	1389	3	US-09-489-039A-2538	

ALIGNMENTS

RESULT 1
US-09-259-214-1
Sequence 1, Application US/09259214A
Patent No. 6110719
GENERAL INFORMATION:
APPLICANT: Kretz, Keith
TITLE OF INVENTION: NOVEL PHYTASE
FILE REFERENCE: DIVER1370-1
CURRENT APPLICATION NUMBER: US/09/259,214A
EARLIER APPLICATION NUMBER: 08/910,798
EARLIER FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1323
TYPE: DNA
ORGANISM: Escherichia coli
FEATURES:
NAME/KEY: CDS
LOCATION: (1)..(1320)
NAME/KEY: misc_feature
LOCATION: (1)..(1323)
OTHER INFORMATION: n = A,T,C or G
US-09-259-214-1

Query Match	100.0%	Score 1322;	DB 3;	Length 1323;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1323;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
1	ATGAAAGGATCTTATCCATTTTATCTCTTGATTCGTTAACCCCGCATCTGCA	60		
1	ATGAAAGGATCTTATCCATTTTATCTCTTGATTCGTTAACCCCGCATCTGCA	60		
61	TTTCGCTCAGAGTGAAGCCGAGAGCTGAAGTGGTGGATTTGCAAGTGGTCTATG	120		
61	TTTCGCTCAGAGTGAAGCCGAGAGCTGAAGTGGTGGATTTGCAAGTGGTCTATG	120		
61	TTTCGCTCAGAGTGAAGCCGAGAGCTGAAGTGGTGGATTTGCAAGTGGTCTATG	120		
121	GTGCGTGTCCAAACCAAGCCAGCACTGATGATGATGATGATGATGATGATGATG	180		
121	GTGCGTGTCCAAACCAAGCCAGCACTGATGATGATGATGATGATGATGATGATG	180		
121	GTGCGTGTCCAAACCAAGCCAGCACTGATGATGATGATGATGATGATGATGATG	180		
181	ACCTGCGCGGTAAACTGGTTGGCTGACACCGCGGTGGTGGATGATGATGATG	240		
181	ACCTGCGCGGTAAACTGGTTGGCTGACACCGCGGTGGTGGATGATGATGATG	240		
181	ACCTGCGCGGTAAACTGGTTGGCTGACACCGCGGTGGTGGATGATGATGATG	240		

QY 241 GGAATTACCAACGCGTAGCGTCTGGTAGCCGACGGAATTGTGCGAATAAAAGGCGTCCG 300
DB 241 GGAATTACCAACGCGTAGCGTCTGGTAGCCGACGGAATTGTGCGAATAAAAGGCGTCCG 300
QY 301 CAGTCTGAGTCAAGTTCGATTAATTCGATGTCGACGAGCGTAACTCCGTAATAACAGCGGAA 360
DB 301 CAGTCTGAGTCAAGTTCGATTAATTCGATGTCGACGAGCGTAACTCCGTAATAACAGCGGAA 360
QY 361 GCCTTCCGCGCGCGGCTGGCACTGACGTGCAATAACCGTACATACCAGCGAGATACG 420
DB 361 GCCTTCCGCGCGCGGCTGGCACTGACGTGCAATAACCGTACATACCAGCGAGATACG 420
QY 421 TCCAGTCCCGATCCGTTATTTAATTCCTTAATAACCTGCGCTTTCGCAACGATTAACGCG 480
DB 421 TCCAGTCCCGATCCGTTATTTAATTCCTTAATAACCTGCGCTTTCGCAACGATTAACGCG 480
QY 481 AACGTGATGACGCGATCTTCAAGCGGCGAGAGGGGTCAATTGCTGACTTTAACGGGCAAT 540
DB 481 AACGTGATGACGCGATCTTCAAGCGGCGAGAGGGGTCAATTGCTGACTTTAACGGGCAAT 540
QY 541 CGGCAAAAGGGGTTTCGGAACTGGAAAGGGGCTTAATTTTCGCAATCAAACTGTGC 600
DB 541 CGGCAAAAGGGGTTTCGGAACTGGAAAGGGGCTTAATTTTCGCAATCAAACTGTGC 600
QY 601 CTTAACCTGAGAAACAGAGCAAAAGCTGTTCAATTAACGACGAGCATTAACATCCGAACTC 660
DB 601 CTTAACCTGAGAAACAGAGCAAAAGCTGTTCAATTAACGACGAGCATTAACATCCGAACTC 660
QY 661 AAGGTAGCGCCGACATGTCTCAATTAACCGGTGCGTAAAGCTTCGACATCAATGCTGACG 720
DB 661 AAGGTAGCGCCGACATGTCTCAATTAACCGGTGCGTAAAGCTTCGACATCAATGCTGACG 720
QY 721 GAGATATTTCTCTGCAACAAGCAAGGAAATGCGGAGCGCGGGTGGGAAAGATCAAC 780
DB 721 GAGATATTTCTCTGCAACAAGCAAGGAAATGCGGAGCGCGGGTGGGAAAGATCAAC 780
QY 781 GATTCAACAACGAGTGAACACCTTGTAAGTTGCAATTAACGCGCAATTTATTTGCTACAA 840
DB 781 GATTCAACAACGAGTGAACACCTTGTAAGTTGCAATTAACGCGCAATTTATTTGCTACAA 840
QY 841 CGCAGCGCAAGAGTTGCCCGGACCGCGCACCCCGTTATTTGATTTGATGATGCGACGCG 900
DB 841 CGCAGCGCAAGAGTTGCCCGGACCGCGCACCCCGTTATTTGATTTGATGATGCGACGCG 900
QY 901 TTGACGCGCCATTCACGCAAAACAGGGGATGAGTGAATTAACCACTTCAGTACTG 960
DB 901 TTGACGCGCCATTCACGCAAAACAGGGGATGAGTGAATTAACCACTTCAGTACTG 960
QY 961 TTATTTGCGGACACGATATCTATCTGCAATCTCGCGCGGCACTGAGCTCAACTG 1020
DB 961 TTATTTGCGGACACGATATCTATCTGCAATCTCGCGCGGCACTGAGCTCAACTG 1020
QY 1021 ACGCTTCCCGGTCAAGCGGATTAACGCGCGGAGTGTGAATCTGTGTTGAACGCTGCG 1080
DB 1021 ACGCTTCCCGGTCAAGCGGATTAACGCGCGGAGTGTGAATCTGTGTTGAACGCTGCG 1080
QY 1081 CGTGGGCTAAGCGATTAACGCGAGTGAATGAGTTTGCTGTGCTTCAGACTTTACG 1140
DB 1081 CGTGGGCTAAGCGATTAACGCGAGTGAATGAGTTTGCTGTGCTTCAGACTTTACG 1140
QY 1141 CAGATGCTGATTAATAACGCGCTGTCAATTAATAACGCGCGCGGAGAGTGAATACTGACC 1200
DB 1141 CAGATGCTGATTAATAACGCGCTGTCAATTAATAACGCGCGCGGAGAGTGAATACTGACC 1200
QY 1201 CTGGCAGGATGTAAGAGCGAAATGCGAGGCGATGTGTTCGTTGAGGTTTAAACGAA 1260
DB 1201 CTGGCAGGATGTAAGAGCGAAATGCGAGGCGATGTGTTCGTTGAGGTTTAAACGAA 1260
QY 1261 ATTCGGAATGAAGACGATACCGGCGTGCAGTTTGAATCTCAATCAACCATCAATAC 1320
DB 1261 ATTCGGAATGAAGACGATACCGGCGTGCAGTTTGAATCTCAATCAACCATCAATAC 1320
QY 1321 TAA 1323

DB 1321 TAA 1323
RESULT 2
US-09-318-528-1
; Sequence 1, Application US/09318528
; Patent No. 6183740
; GENERAL INFORMATION:
; APPLICANT: Kreitz, Kelch
; TITLE OF INVENTION: NOVEL PHRYASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/318,528
; EARLIER FILING DATE: 1999-05-25
; EARLIER APPLICATION NUMBER: 09/291,931
; EARLIER FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1320)
; NAME/KEY: misc feature
; LOCATION: (1)..(1323)
; OTHER INFORMATION: n = A,T,C or G
US-09-318-528-1
Query Match 100.0%; Score 1322; DB 3; Length 1323;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAAGCATCTTAATCCCATTTTATCTTGTGATTCGATTAACCGGCAATTCGCA 60
DB 1 ATGAAAGCATCTTAATCCCATTTTATCTTGTGATTCGATTAACCGGCAATTCGCA 60
QY 61 TTCCTGAGAGTGAAGCGGAGCTGAAAGTGTGATTCAGTCTCATGCT 120
DB 61 TTCCTGAGAGTGAAGCGGAGCTGAAAGTGTGATTCAGTCTCATGCT 120
QY 121 GTGCTGCTCCAAACCAAGGCCCAAGCACTGATGCAAGATTCACCCAGACGATGGCCA 180
DB 121 GTGCTGCTCCAAACCAAGGCCCAAGCACTGATGCAAGATTCACCCAGACGATGGCCA 180
QY 181 ACCTGGCCGGTAAACTGGGTGCTGACACCGCGAGGTGAGCTTAATGCGCTATCTC 240
DB 181 ACCTGGCCGGTAAACTGGGTGCTGACACCGCGAGGTGAGCTTAATGCGCTATCTC 240
QY 241 GGAATTACCAACGCGCTCTGCTGAGCGGACGATTTGCTGCGAATAAAAGGCGTCCG 300
DB 241 GGAATTACCAACGCGCTCTGCTGAGCGGACGATTTGCTGCGAATAAAAGGCGTCCG 300
QY 301 CAGTCTGAGTCAAGTTCGATTAATTCGATGTCGACGAGCGTAACTCCGTAATAACAGCGGAA 360
DB 301 CAGTCTGAGTCAAGTTCGATTAATTCGATGTCGACGAGCGTAACTCCGTAATAACAGCGGAA 360
QY 361 GCCTTCCGCGCGGCTGGCACTGACGTGCAATAACCGTACATACCAGCGAGATACG 420
DB 361 GCCTTCCGCGCGGCTGGCACTGACGTGCAATAACCGTACATACCAGCGAGATACG 420
QY 421 TCCAGTCCCGATCCGTTATTTAATTCCTTAATAACCTGCGCTTTCGCAACGATTAACGCG 480
DB 421 TCCAGTCCCGATCCGTTATTTAATTCCTTAATAACCTGCGCTTTCGCAACGATTAACGCG 480
QY 481 AACGTGATGACGCGATCTTCAAGCGGCGAGAGGGGTCAATTGCTGACTTTAACGGGCAAT 540
DB 481 AACGTGATGACGCGATCTTCAAGCGGCGAGAGGGGTCAATTGCTGACTTTAACGGGCAAT 540

```

QY 541 CGGCAACGGCGTTTGGCGAAGTGAACGGGCTTAATTTCCGCAATCAAACTTGTGC 600
   |||
Db 541 CGGCAACGGCGTTTGGCGAAGTGAACGGGCTTAATTTCCGCAATCAAACTTGTGC 600
QY 601 CTTAAACGTGAAACAGACAGAAAGCTGTTCAATTAACGACGATTAACATCGAATC 660
   |||
Db 601 CTTAAACGTGAAACAGACAGAAAGCTGTTCAATTAACGACGATTAACATCGAATC 660
QY 661 AAGGTGAGCGCGAACAATGTTCTCATTAACCGGTGGGTAAAGCTCGCATCAATGCTGACG 720
   |||
Db 661 AAGGTGAGCGCGAACAATGTTCTCATTAACCGGTGGGTAAAGCTCGCATCAATGCTGACG 720
QY 721 GAGATATTTCTCTGTGAACAAGACAGGAAATGCGGACCGGGGTGGGAAAGATCACC 780
   |||
Db 721 GAGATATTTCTCTGTGAACAAGACAGGAAATGCGGACCGGGGTGGGAAAGATCACC 780
QY 781 GATTCAACACAGTGAACAACCTTGTCAAGTTTGATTAAGCGCAATTTTATTGCTACAA 840
   |||
Db 781 GATTCAACACAGTGAACAACCTTGTCAAGTTTGATTAAGCGCAATTTTATTGCTACAA 840
QY 841 CGCAGCGCAGAGGTTGCGCGACGCGCCACCCCGTTATTGATTGATTCATGGCAGCG 900
   |||
Db 841 CGCAGCGCAGAGGTTGCGCGACGCGCCACCCCGTTATTGATTGATTCATGGCAGCG 900
QY 901 TTGAGCGCCCATCCACCGGAAAAAGGGGTATGATGATCAATTAACCTTCAATGATC 960
   |||
Db 901 TTGAGCGCCCATCCACCGGAAAAAGGGGTATGATGATCAATTAACCTTCAATGATC 960
QY 961 TTTATTTGCGGCAACAAGTACTAATCTGGAATCTGCGCGGCGCACTGGAAGCTCAACTG 1020
   |||
Db 961 TTTATTTGCGGCAACAAGTACTAATCTGGAATCTGCGCGGCGCACTGGAAGCTCAACTG 1020
QY 1021 ACGCTTCCCGGTCAAGCGGATTAACGCGCGCAAGGTGTAACCTGTTTGAACGCTG 1080
   |||
Db 1021 ACGCTTCCCGGTCAAGCGGATTAACGCGCGCAAGGTGTAACCTGTTTGAACGCTG 1080
QY 1081 CGTCCGCTTAAGGATTAACGCGGATGATTCAGGTTTCCGTCTTCAAGCTTTACG 1140
   |||
Db 1081 CGTCCGCTTAAGGATTAACGCGGATGATTCAGGTTTCCGTCTTCAAGCTTTACG 1140
QY 1141 CAGATGCGTGAATAACGCGCGCTGTCTAATAATAGCGCGCGCGGAGGTTGAACCTGAC 1200
   |||
Db 1141 CAGATGCGTGAATAACGCGCGCTGTCTAATAATAGCGCGCGCGGAGGTTGAACCTGAC 1200
QY 1201 CTGGCAGATGTGAAGAGGAAATGCGCAGGCGATGTGTTGTCGTCAGGTTTACGCA 1260
   |||
Db 1201 CTGGCAGATGTGAAGAGGAAATGCGCAGGCGATGTGTTGTCGTCAGGTTTACGCA 1260
QY 1261 ATCGTGAATGAAGACGATACCGGCGTGCAGTTTGAATCTCATCAATCAATCAG 1320
   |||
Db 1261 ATCGTGAATGAAGACGATACCGGCGTGCAGTTTGAATCTCATCAATCAG 1320
QY 1321 TAA 1323
   |||
Db 1321 TAA 1323

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; NAME/KEY: misc_feature
; LOCATION: (1)...(1323)
; OTHER INFORMATION: n = A,T,C or G
US-09-291-931-1

Query Match      100.0%; Score 1322; DB 3; Length 1323;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGGATCTTATCCATTTTATCTCTTGATTCGGTTAACCCGCAATCTGCA 60
   |||
Db 1 ATGAAGGATCTTATCCATTTTATCTCTTGATTCGGTTAACCCGCAATCTGCA 60
QY 61 TTGCTCAGAGTGAACCGGAGCTGAAGCTGAAGTGTGTGATTGTCAGTGCTAGGT 120
   |||
Db 61 TTGCTCAGAGTGAACCGGAGCTGAAGCTGAAGTGTGTGATTGTCAGTGCTAGGT 120
QY 121 GTGCTGCTCCACCAAGGCGACGCACTGATGACGATGTCAATCCCGACGATGGCCA 180
   |||
Db 121 GTGCTGCTCCACCAAGGCGACGCACTGATGACGATGTCAATCCCGACGATGGCCA 180
QY 181 ACTCGCGGTTAAACTGGGTTGGCTGACACCGGAGGAGTGAAGTATGCTATCTC 240
   |||
Db 181 ACTCGCGGTTAAACTGGGTTGGCTGACACCGGAGGAGTGAAGTATGCTATCTC 240
QY 241 GGAATTAACAACGCGCAGCGTCTGTGAGCCGACGATTTGCTGGCCGAAAAAGGCTGCGC 300
   |||
Db 241 GGAATTAACAACGCGCAGCGTCTGTGAGCCGACGATTTGCTGGCCGAAAAAGGCTGCGC 300
QY 301 CAGTCTGTCAGGTGGCGATTATGCTGATGTCGACGAGCTTACCCGTTAAACAGGCGAA 360
   |||
Db 301 CAGTCTGTCAGGTGGCGATTATGCTGATGTCGACGAGCTTACCCGTTAAACAGGCGAA 360
QY 361 GCTTGGCGCGGCGGCTGACCTGACCTGTGCAATTAACGATTAACCAAGGCAATACG 420
   |||
Db 361 GCTTGGCGCGGCGGCTGACCTGACCTGTGCAATTAACGATTAACCAAGGCAATACG 420
QY 421 TCCAGTCCCGATCCGTTATTTAATCTCTTAAAACTGGCGTTTGGCAATGATTAACGCG 480
   |||
Db 421 TCCAGTCCCGATCCGTTATTTAATCTCTTAAAACTGGCGTTTGGCAATGATTAACGCG 480
QY 481 AAGTGAATGACGCGATCTTCAAGCAGGCGAGAGGCTCAATTTGCTGACTTTACCGGGCAT 540
   |||
Db 481 AAGTGAATGACGCGATCTTCAAGCAGGCGAGAGGCTCAATTTGCTGACTTTACCGGGCAT 540
QY 541 CGGCAACGGCGTTTGGGAACTGGAAGGGGCTTAATTTCCGCAATCAAACTTGTGC 600
   |||
Db 541 CGGCAACGGCGTTTGGGAACTGGAAGGGGCTTAATTTCCGCAATCAAACTTGTGC 600
QY 601 CTTAAACGTGAAACAGACAGAAAGCTGTTCAATTAACGACGATTAACATCGAATC 660
   |||
Db 601 CTTAAACGTGAAACAGACAGAAAGCTGTTCAATTAACGACGATTAACATCGAATC 660
QY 661 AAGGTGAGCGCGAACAATGTTCTCATTAACCGGTGGGTAAAGCTCGCATCAATGCTGACG 720
   |||
Db 661 AAGGTGAGCGCGAACAATGTTCTCATTAACCGGTGGGTAAAGCTCGCATCAATGCTGACG 720
QY 721 GAGATATTTCTCTGTGAACAAGACAGGAAATGCGGACCGGGGTGGGAAAGATCACC 780
   |||
Db 721 GAGATATTTCTCTGTGAACAAGACAGGAAATGCGGACCGGGGTGGGAAAGATCACC 780
QY 781 GATTCAACACAGTGAACAACCTTGTCAAGTTTGATTAAGCGCAATTTTATTGCTACAA 840
   |||
Db 781 GATTCAACACAGTGAACAACCTTGTCAAGTTTGATTAAGCGCAATTTTATTGCTACAA 840

```

QY 841 CGCAGCGAGAGTGGTCCCGGAGCCGCGCACCCCGTATTGATTTGATCATGCGACG 900
| | | | |
Db 841 CGCAGCGAGAGTGGTCCCGGAGCCGCGCACCCCGTATTGATTTGATCATGCGACG 900
QY 901 TTGACGCCCATTCACCGGCAAAAACAGGCGTATGATGATCACTTACCTTCACTG 960
| | | | |
Db 901 TTGACGCCCATTCACCGGCAAAAACAGGCGTATGATGATCACTTACCTTCACTG 960
QY 961 TTTATTTGCGGAGACAGATACTTAATCTGGCAAAATCTCGGCGGCGCATCTGAGCTCAACTG 1020
| | | | |
Db 961 TTTATTTGCGGAGACAGATACTTAATCTGGCAAAATCTCGGCGGCGCATCTGAGCTCAACTG 1020
QY 1021 ACGCTTCCCGGTGACGCGGATTAACAAGCGCGCATGATGTAATCTGTTTGAACGCTG 1080
| | | | |
Db 1021 ACGCTTCCCGGTGACGCGGATTAACAAGCGCGCATGATGTAATCTGTTTGAACGCTG 1080
QY 1081 CGTCGGCTTAAGCGATTAACAGCAGTGAATTCAGGTTTCTGCTGCTTCCAGACTTTACAG 1140
| | | | |
Db 1081 CGTCGGCTTAAGCGATTAACAGCAGTGAATTCAGGTTTCTGCTGCTTCCAGACTTTACAG 1140
QY 1141 CAGATGCGGTATTAACAGCGCGCTGCTATTAATATGCGCGCGCGGAGAGTGAACCTGACC 1200
| | | | |
Db 1141 CAGATGCGGTATTAACAGCGCGCTGCTATTAATATGCGCGCGCGGAGAGTGAACCTGACC 1200
QY 1201 CTGGCAGGATGTAAGAGCGAATGCGAGGCGCATGTTGTTGCGCAGTTTTCAGCA 1260
| | | | |
Db 1201 CTGGCAGGATGTAAGAGCGAATGCGAGGCGCATGTTGTTGCGCAGTTTTCAGCA 1260
QY 1261 ATCGTGAATGAAGCAGCATACCGGCGTGCAGTTTGAGATCTCATCACCATCATTAC 1320
| | | | |
Db 1261 ATCGTGAATGAAGCAGCATACCGGCGTGCAGTTTGAGATCTCATCACCATCATTAC 1320
QY 1321 TAA 1323
| | | | |
Db 1321 TAA 1323
| | | | |

RESULT 4

US-09-580-515-1

Sequence 1, Application US/09580515
Patent No. 6720014
GENERAL INFORMATION:
APPLICANT: Kretz, Keith
TITLE OF INVENTION: NOVEL PHYTASE
FILE REFERENCE: 09010/029003
CURRENT APPLICATION NUMBER: US/09/580,515
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 09/318,528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: 08/910,798
PRIOR FILING DATE: 1997-08-13
PRIOR APPLICATION NUMBER: 09/259,214
PRIOR FILING DATE: 1999-03-01
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1323
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1320)
NAME/KEY: misc_feature
LOCATION: (1)...(1323)
OTHER INFORMATION: n = A,T,C or G
US-09-580-515-1

Query Match 100.0%; Score 1322; DB 3; Length 1323;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAGCATCTTAATCCATTTTATCTCTTGATTCGGTTAAACCCCGCAATCTGCA 60
| | | | |
Db 1 ATGAAAGCATCTTAATCCATTTTATCTCTTGATTCGGTTAAACCCCGCAATCTGCA 60
QY 61 TTGCTCAAGTGAAGCCGAGCTGAGCTGAAAGTGTGATTTGTCAGTGTGATGCT 120
| | | | |
Db 61 TTGCTCAAGTGAAGCCGAGCTGAGCTGAAAGTGTGATTTGTCAGTGTGATGCT 120
QY 121 GTGGTGTCTCCAAACCAAGGCGGAGCACTGATGAGATGCAACCCGAGACGATGCGCA 180
| | | | |
Db 121 GTGGTGTCTCCAAACCAAGGCGGAGCACTGATGAGATGCAACCCGAGACGATGCGCA 180
QY 181 ACTGCGCGGTAAACTGAGTTGCTGCAACCGCGAGTGTGAGTAAATCGCCTATCTC 240
| | | | |
Db 181 ACTGCGCGGTAAACTGAGTTGCTGCAACCGCGAGTGTGAGTAAATCGCCTATCTC 240
QY 241 GGAATTAACAAAGCCAGCGCTGAGTACCGAATTCCTGCGGAAAAGGCTGCCCG 300
| | | | |
Db 241 GGAATTAACAAAGCCAGCGCTGAGTACCGAATTCCTGCGGAAAAGGCTGCCCG 300
QY 301 CAGTCTGCTCAAGTCCGAGTTATTTGCTGATGTCAGACGAGGTACCCGTTAAACAGCGCAA 360
| | | | |
Db 301 CAGTCTGCTCAAGTCCGAGTTATTTGCTGATGTCAGACGAGGTACCCGTTAAACAGCGCAA 360
QY 361 GCTTTCGCGCGCGGCTGCGCACTGACTGTGCAATTAACCGTATACCCAGCGAGATACG 420
| | | | |
Db 361 GCTTTCGCGCGCGGCTGCGCACTGACTGTGCAATTAACCGTATACCCAGCGAGATACG 420
QY 421 TCCAGTCCCGATCCGTTATTTAATCTCTTAAATCTGGCGTTTCCCACTGAGTAAGCGG 480
| | | | |
Db 421 TCCAGTCCCGATCCGTTATTTAATCTCTTAAATCTGGCGTTTCCCACTGAGTAAGCGG 480
QY 481 AAGCTGACTGACGCGATCTCAGACGAGGAGGATCAATGCTGACTTAAACCGGCGAT 540
| | | | |
Db 481 AAGCTGACTGACGCGATCTCAGACGAGGAGGATCAATGCTGACTTAAACCGGCGAT 540
QY 541 CGGCAAAAGCGGTTTCGCGCACTGGAACGAGTGTCTTAATTTCCGCAATCAACTGTGC 600
| | | | |
Db 541 CGGCAAAAGCGGTTTCGCGCACTGGAACGAGTGTCTTAATTTCCGCAATCAACTGTGC 600
QY 601 CTTAAACGTGAAACAGAGCAAAAGCTGTTCAATTAACGAGCATTAACATCGAAATCTC 660
| | | | |
Db 601 CTTAAACGTGAAACAGAGCAAAAGCTGTTCAATTAACGAGCATTAACATCGAAATCTC 660
QY 661 AAGGTGAGCGCGGCAATGCTCTTAACCGGTGCGTAAAGCTCGATCAATGCTGACG 720
| | | | |
Db 661 AAGGTGAGCGCGGCAATGCTCTTAACCGGTGCGTAAAGCTCGATCAATGCTGACG 720
QY 721 GAGATATTTCTCTGCAACAAAGCAGGGAATGCGGAGCGGAGTGGGAGATCAAC 780
| | | | |
Db 721 GAGATATTTCTCTGCAACAAAGCAGGGAATGCGGAGCGGAGTGGGAGATCAAC 780
QY 781 GATTCAACACAGTGAACACCTTGTCTAAGTTTGCAATACGCGCAATTTTATTTGCTACAA 840
| | | | |
Db 781 GATTCAACACAGTGAACACCTTGTCTAAGTTTGCAATACGCGCAATTTTATTTGCTACAA 840
QY 841 GATTCAACACAGTGAACACCTTGTCTAAGTTTGCAATACGCGCAATTTTATTTGCTACAA 840
| | | | |
Db 841 GATTCAACACAGTGAACACCTTGTCTAAGTTTGCAATACGCGCAATTTTATTTGCTACAA 840
QY 841 CGCAGCGCAGAGGTTGCCCGGAGCGCGCAACCCGTTATTTGATTTGATGATGCGGCG 900
| | | | |
Db 841 CGCAGCGCAGAGGTTGCCCGGAGCGCGCAACCCGTTATTTGATTTGATGATGCGGCG 900
QY 901 TTGACGCCCATTCACCGGCAAAAACAGGCGTATGATGATCACTTACCTTCACTG 960
| | | | |
Db 901 TTGACGCCCATTCACCGGCAAAAACAGGCGTATGATGATCACTTACCTTCACTG 960
QY 961 TTTATTTGCGGAGACAGATACTTAATCTGGCAAAATCTCGGCGGCGCATCTGAGCTCAACTG 1020
| | | | |
Db 961 TTTATTTGCGGAGACAGATACTTAATCTGGCAAAATCTCGGCGGCGCATCTGAGCTCAACTG 1020
QY 1021 ACGCTTCCCGGTGACGCGGATTAACAAGCGCGCATGATGTAATCTGTTTGAACGCTG 1080
| | | | |
Db 1021 ACGCTTCCCGGTGACGCGGATTAACAAGCGCGCATGATGTAATCTGTTTGAACGCTG 1080
QY 1081 CGTCGGCTTAAGCGATTAACAGCAGTGAATTCAGGTTTCTGCTGCTTCCAGACTTTACAG 1140
| | | | |

Db 1081 CGTCGGCTAAGCATTAACGCGCAGTGGATTTCAGGTTTCGCTGATCTTCCAGACTTTACAG 1140
Qy 1141 CAGATCCGCTGATTAACGCCCGCTGTCATTAAATACGCCCGCGAGAGGTGAACCTGACC 1200
Db 1141 CAGATCCGCTGATTAACGCCCGCTGTCATTAAATACGCCCGCGAGAGGTGAACCTGACC 1200
Qy 1201 CTGGCAGATGTGAAGAGGAAATGCGCAGGCGCATGTGTTGTCAGAGTTTACGCA 1260
Db 1201 CTGGCAGATGTGAAGAGGAAATGCGCAGGCGCATGTGTTGTCAGAGTTTACGCA 1260
Qy 1261 ATCGTAATGAAGCAGCATACCGCGCTGCAAGTTGAGATCTCATACCATCATAC 1320
Db 1261 ATCGTAATGAAGCAGCATACCGCGCTGCAAGTTGAGATCTCATACCATCATAC 1320
Qy 1321 TAA 1323
Db 1321 TAA 1323

RESULT 5

US-09-866-379-1

Sequence 1, Application US/09866379

Patent No. 6855365

GENERAL INFORMATION:

APPLICANT: DIVERSA CORPORATION

APPLICANT: SHORT, Jay

APPLICANT: KREIZ, Keith

APPLICANT: GRAY, Kevin

APPLICANT: BARTON, Nelson

APPLICANT: GARRETT, James

APPLICANT: O'DONOGHUE, Kileen

TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF

FILE REFERENCE: DIVER1370-7

CURRENT APPLICATION NUMBER: US/09/866,379

CURRENT FILING DATE: 2001-05-24

PRIOR APPLICATION NUMBER: US 09/580,515

PRIOR FILING DATE: 2000-05-25

PRIOR APPLICATION NUMBER: US 09/318,528

PRIOR FILING DATE: 1999-05-25

PRIOR APPLICATION NUMBER: US 09/291,931

PRIOR FILING DATE: 1999-04-13

PRIOR APPLICATION NUMBER: US 09/259,214

PRIOR FILING DATE: 1999-03-01

PRIOR APPLICATION NUMBER: US 08/910,798

PRIOR FILING DATE: 1997-08-13

NUMBER OF SEQ ID NOS: 10

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 1323

TYPE: DNA

ORGANISM: Escherichia coli

FEATURE:

NAME/KEY: misc feature

LOCATION: (1) .. (1323)

OTHER INFORMATION: n is any nucleotide

NAME/KEY: CDS

LOCATION: (1) .. (1323)

OTHER INFORMATION:

US-09-866-379-1

Query Match 100.0%; Score 1322; DB 3; Length 1323;
Best local similarity 100.0%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAGGAGATCTTATCCCATTTTATCTCTTGTATCCGTTAACCCCGCATCTGCA 60
Db 1 ATGAAGGAGATCTTATCCCATTTTATCTCTTGTATCCGTTAACCCCGCATCTGCA 60
Qy 61 TTGCTCAGAGTGAAGCCGAGCTGAAGCTGAAAGTGTGTGATTTGTCAGTGTGATGAT 120
Db 61 TTGCTCAGAGTGAAGCCGAGCTGAAGCTGAAAGTGTGTGATTTGTCAGTGTGATGAT 120

Qy 121 GTGCGGTCTCAACCAAGGCCACGCAACTGATGCAAGATGTCAACCCCAAGCGATGCGCA 180
Db 121 GTGCGGTCTCAACCAAGGCCACGCAACTGATGCAAGATGTCAACCCCAAGCGATGCGCA 180
Qy 181 ACCGTGCGGTGAAAATGAGGTTTGCTGACACCGCGAGGTGTGAGTAAATGCTATCTC 240
Db 181 ACCGTGCGGTGAAAATGAGGTTTGCTGACACCGCGAGGTGTGAGTAAATGCTATCTC 240
Qy 241 GGAATTACCAACGCCAGGCTGTGATGCGCAAGATTGCTGCGCAAAAAGAGGCTGCGC 300
Db 241 GGAATTACCAACGCCAGGCTGTGATGCGCAAGATTGCTGCGCAAAAAGAGGCTGCGC 300
Qy 301 CAGTCTGCTCAAGTGGCGATTTATGCTGATGTGCAAGACGCTAACCCGTAAACAGCGCA 360
Db 301 CAGTCTGCTCAAGTGGCGATTTATGCTGATGTGCAAGACGCTAACCCGTAAACAGCGCA 360
Qy 361 GCCTTGCGCGGCTGCGCACTGACCTGCAATGCAATACCTGATCCAGGCGATGAC 420
Db 361 GCCTTGCGCGGCTGCGCACTGACCTGCAATGCAATACCTGATCCAGGCGATGAC 420
Qy 421 TCCAGTCCCGATCCGTTATTTATCTCTTAAACCTGAGGCTTGCCTGCAATGACGCG 480
Db 421 TCCAGTCCCGATCCGTTATTTATCTCTTAAACCTGAGGCTTGCCTGCAATGACGCG 480
Qy 481 AACGTGACTGAAGCGCATCTCAGCAGGCGAGAGGCTCAATTGCTGACTTTACCGGCGAT 540
Db 481 AACGTGACTGAAGCGCATCTCAGCAGGCGAGAGGCTCAATTGCTGACTTTACCGGCGAT 540
Qy 541 CGGCAAAAGGCGGCTTGCAGAACTGGAACGGGCTTAAATTTTCCGCAATCAACTGTGCG 600
Db 541 CGGCAAAAGGCGGCTTGCAGAACTGGAACGGGCTTAAATTTTCCGCAATCAACTGTGCG 600
Qy 601 CTTAAACGTGAGAAACAGAACGAAAGCTGTTCAATTAACGACGCAATTAACATCGAATCTC 660
Db 601 CTTAAACGTGAGAAACAGAACGAAAGCTGTTCAATTAACGACGCAATTAACATCGAATCTC 660
Qy 661 AAGGTGAGCGCGCAAAATGTCTCAATTAACCGGTGCGTAAAGCTGCGATCAATGCTGAGC 720
Db 661 AAGGTGAGCGCGCAAAATGTCTCAATTAACCGGTGCGTAAAGCTGCGATCAATGCTGAGC 720
Qy 721 GAGATATTTCTCTGCAACAAAGCAAGGAAATGCGGAGCGGAGGAGGAAAGATCAAC 780
Db 721 GAGATATTTCTCTGCAACAAAGCAAGGAAATGCGGAGCGGAGGAGGAAAGATCAAC 780
Qy 781 GATTCAACCAAGTGAACACCTTGTGAATTTGCAATTAACGCGCAATTTTATTTGCTACAA 840
Db 781 GATTCAACCAAGTGAACACCTTGTGAATTTGCAATTAACGCGCAATTTTATTTGCTACAA 840
Qy 841 CGCAGCGCAAGGTTGCGCGCAGCGCGCACCCCGTTATGATTTGATCATGCGACGCG 900
Db 841 CGCAGCGCAAGGTTGCGCGCAGCGCGCACCCCGTTATGATTTGATCATGCGACGCG 900
Qy 901 TTGACGCCCATTCACCGCAAAAACAGGCGTATGCTGACATTAACCATTCAGTACG 960
Db 901 TTGACGCCCATTCACCGCAAAAACAGGCGTATGCTGACATTAACCATTCAGTACG 960
Qy 961 TTTATTTGCGGACAGATACTAATCTGCAATCTGCGCGCGCACTGAGAGCTCAACTGCG 1020
Db 961 TTTATTTGCGGACAGATACTAATCTGCAATCTGCGCGCGCACTGAGAGCTCAACTGCG 1020
Qy 1021 AAGCTTCCCGGTGACGCGGATTAACCGCGCGCAAGTGTGTAATCTGTTGAAAGCTGCG 1080
Db 1021 AAGCTTCCCGGTGACGCGGATTAACCGCGCGCAAGTGTGTAATCTGTTGAAAGCTGCG 1080
Qy 1081 CGTGGCTAAGCAGTAAACAGCAGGATTAACGAGTTTGCCTGCTTCCAGACTTAAAG 1140
Db 1081 CGTGGCTAAGCAGTAAACAGCAGGATTAACGAGTTTGCCTGCTTCCAGACTTAAAG 1140
Qy 1141 CAGATGCGTGAATTAACCGCGCTGTCAATTAATACCGCGCGAGAGGTGAACCTGACC 1200
Db 1141 CAGATGCGTGAATTAACCGCGCTGTCAATTAATACCGCGCGAGAGGTGAACCTGACC 1200
Qy 1201 CTGGCAGATGTGAAGAGGAAATGCGCAGGCGCATGTGTTGTCAGAGTTTACGCA 1260

Db 1201 CTGGCAGGATGTGAAGACGCAATGCGAGGCGATGTGTTGCGAGGTTTACGCA 1260
Qy 1261 ATCGTAATGAAGACGCAATCCGCGTGCAGTTTGATGATCTCATCACCATCACCATCAC 1320
Db 1261 ATCGTAATGAAGACGCAATCCGCGCGTGCAGTTTGATGATCTCATCACCATCACCATCAC 1320
Qy 1321 TAA 1323
Db 1321 TAA 1323

RESULT 6

US-09-866-379-7
Sequence 7, Application US/09866379
Patent No. 6855365
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: SHORT, Jay
APPLICANT: KRETZ, Keith
APPLICANT: GRAY, Kevin
APPLICANT: BARTON, Nelson
APPLICANT: GARRETT, James
APPLICANT: O'DONOGHUE, Bileen
TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
FILE REFERENCE: DIVER1370-7
CURRENT APPLICATION NUMBER: US/09/866,379
PRIOR FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: US 09/580,515
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 09/318,528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: US 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 09/259,214
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: US 08/910,798
PRIOR FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7
LENGTH: 1901
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(1901)
OTHER INFORMATION: n is any nucleotide
US-09-866-379-7

Query Match 97.4%; Score 1287.8; DB 3; Length 1901;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1293; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ATGAAGGATCTTAATCCATTTTATCTCTGATTCGGTTAAACCCCGCATCTGCA 60
Db 188 ATGAAGGATCTTAATCCATTTTATCTCTGATTCGGTTAAACCCCGCATCTGCA 247
Qy 61 TTGCTCAGAGTGAGCCGGAAGCTGGAAGTGTGATTTGTCAGTCTGATGAT 120
Db 248 TTGCTCAGAGTGAGCCGGAAGCTGGAAGTGTGATTTGTCAGTCTGATGAT 307
Qy 121 GTGCGTCTCCAAACCAAGCCGCACTGATGCAAGATGTCACCCGAGCATGAGCA 180
Db 308 GTGCGTCTCCAAACCAAGCCGCACTGATGCAAGATGTCACCCGAGCATGAGCA 367
Qy 181 ACCTGGCCGGTAAACCTGGTTGGCTGACACCGCGGGTGGAGCTAATCGCTATCTC 240
Db 368 ACCTGGCCGGTAAACCTGGTTGGCTGACACCGCGGGTGGAGCTAATCGCTATCTC 427
Qy 241 GGAATTACCAACGCAAGCTCTGCTAGCCGACGAGATTGTCGCAAAAAAGGCTGCCG 300
Db 428 GGAATTACCAACGCAAGCTCTGCTAGCCGACGAGATTGTCGCAAAAAAGGCTGCCG 487

Qy 301 CAGTCTGCTCAGGTCGGGATTTATGCTGATGTGCAAGAGGTACCGGTTAAACAGCGAA 360
Db 488 CAGTCTGCTCAGGTCGGGATTTATGCTGATGTGCAAGAGGTACCGGTTAAACAGCGAA 547
Qy 361 GCCTTCGCCCGCGGGCTGCGCACTGACTGTCATAACCGTACATACCCAGGAGATACG 420
Db 548 GCCTTCGCCCGCGGGCTGCGCACTGACTGTCATAACCGTACATACCCAGGAGATACG 607
Qy 421 TCCAGTCCCGATCCGTTATTTATCTCTAATAAACCTGCGTTTGCCAACTGGATACGCG 480
Db 608 TCCAGTCCCGATCCGTTATTTATCTCTAATAAACCTGCGTTTGCCAACTGGATACGCG 667
Qy 481 AACGTGACTAGCGGATCTCTCAGAGGCGAGAGGTCATGTCGACTTAAACGGGCAT 540
Db 668 AACGTGACTAGCGGATCTCTCAGAGGCGAGAGGTCATGTCGACTTAAACGGGCAT 727
Qy 541 CGGCAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCGCAATCMAACTGTGC 600
Db 728 CGGCAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCGCAATCMAACTGTGC 787
Qy 601 CTTAACGTGAGAAACAGAACGAAAGCTGTTCAATTAAACGACGATTAACATCGAACTC 660
Db 788 CTTAACGTGAGAAACAGAACGAAAGCTGTTCAATTAAACGACGATTAACATCGAACTC 847
Qy 661 AAGGTGAGCGCGCCGACATGTCCTAATAACCGGTGCGGTAAAGCTCGCATCAATGCTGACG 720
Db 848 AAGGTGAGCGCGCCGACATGTCCTAATAACCGGTGCGGTAAAGCTCGCATCAATGCTGACG 907
Qy 721 GAGATATTTCTCTGCAACAGCAAGCAAGATGCGAGCGCGGGTGGAGAGATCAC 780
Db 908 GAGATATTTCTCTGCAACAGCAAGCAAGATGCGAGCGCGGGTGGAGAGATCAC 967
Qy 781 GATTCAACACAGTGAACACCTTGCTAAGTTTGATTAACGCGCAATTTTATTTGCTACAA 840
Db 968 GATTCAACACAGTGAACACCTTGCTAAGTTTGATTAACGCGCAATTTTATTTGCTACAA 1027
Qy 841 CGCACGCGCAAGGTGCGCGGAGCGCGGCAACCCCGTTATTTGATTTGATCATGCGCGG 900
Db 1028 CGCACGCGCAAGGTGCGCGGAGCGCGGCAACCCCGTTATTTGATTTGATCATGCGCGG 1087
Qy 901 TTGACGCCCATCCACCGCAAAACAGAGCGATGCTGATGACATTAACCATTCAGTACTG 960
Db 1088 TTGACGCCCATCCACCGCAAAACAGAGCGATGCTGATGACATTAACCATTCAGTACTG 1147
Qy 961 TTATATCCCGGACAGATTAATCTG3CAAAATCTG3CGCGGCACTGAGCTCACTGG 1020
Db 1148 TTATATCCCGGACAGATTAATCTG3CAAAATCTG3CGCGGCACTGAGCTCACTGG 1207
Qy 1021 ACGCTTCCCGGTGACGCGGATTAACAGCGCGGCAAGTGTGAACCTG3GTTGAACGCGTGG 1080
Db 1208 ACGCTTCCCGGTGACGCGGATTAACAGCGCGGCAAGTGTGAACCTG3GTTGAACGCGTGG 1267
Qy 1081 CGTGGCTAAGCGATTAACAGCGAGTGTGATTCAGGTTTCGCTGCTTCAGACTTTACAG 1140
Db 1268 CGTGGCTAAGCGATTAACAGCGAGTGTGATTCAGGTTTCGCTGCTTCAGACTTTACAG 1327
Qy 1141 CAGATGCGTGAATAAACGCGCTGTCAATTAAATACGCGCCCGGAGAGGTGAACCTGAC 1200
Db 1328 CAGATGCGTGAATAAACGCGCTGTCAATTAAATACGCGCCCGGAGAGGTGAACCTGAC 1387
Qy 1201 CTGGCAGAGTGTGAAGCGCAATGCGCAGGCGCATGTTGCTGTGGAGAGTTTAAACGCA 1260
Db 1388 CTGGCAGAGTGTGAAGCGCAATGCGCAGGCGCATGTTGCTGTGGAGAGTTTAAACGCA 1447
Qy 1261 ATCGTAATGAAGACGCAATACCGGCGTGCAGTTTGAGAT 1300
Db 1448 ATCGTAATGAAGACGCAATACCGGCGTGCAGTTTGAT 1487

RESULT 7

US-09-866-379-5
Sequence 5, Application US/09866379

```
/ Patent No. 6855365
/ GENERAL INFORMATION:
/ APPLICANT: DIVERSA CORPORATION
/ APPLICANT: SHORT, Jay
/ APPLICANT: KREIZ, Keith
/ APPLICANT: GRAY, Kevin
/ APPLICANT: BARTON, Nelson
/ APPLICANT: GABRETT, James
/ APPLICANT: O'DONOHUE, Bileen
/ TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
/ FILE REFERENCE: DIVER1370-7
/ CURRENT APPLICATION NUMBER: US/09/866,379
/ PRIOR APPLICATION NUMBER: US 09/580,515
/ PRIOR FILING DATE: 2000-05-25
/ PRIOR APPLICATION NUMBER: US 09/318,528
/ PRIOR FILING DATE: 1999-05-25
/ PRIOR APPLICATION NUMBER: US 09/291,931
/ PRIOR FILING DATE: 1999-04-13
/ PRIOR APPLICATION NUMBER: US 09/259,214
/ PRIOR FILING DATE: 1999-03-01
/ PRIOR APPLICATION NUMBER: US 08/910,798
/ PRIOR FILING DATE: 1997-08-13
/ NUMBER OF SEQ ID NOS: 10
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 5
/ LENGTH: 1901
/ TYPE: DNA
/ ORGANISM: Escherichia coli
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)-(1901)
/ OTHER INFORMATION: n is any nucleotide
US-09-866-379-5
```

```
Query Match      97.3%; Score 1286.2; DB 3; Length 1901;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1992; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
QY      1 ATGAAGCGATCTTAATCCCAATTTTATCTCTTGATTCGGTTAACCCCGCAATCTGCA 60
DB      188 ATGAAGCGATCTTAATCCCAATTTTATCTCTTGATTCGGTTAACCCCGCAATCTGCA 247
QY      61 TTGCGTCAGAGTGAAGCCGGAAGCTGAAGCTGAAGTGTGTGATTTCTAGTGTATGAT 120
DB      248 TTGCGTCAGAGTGAAGCCGGAAGCTGAAGCTGAAGTGTGTGATTTCTAGTGTATGAT 307
QY      121 GTGCGTGTCCAACCAAGGCGCAAGCACTGATGCAAGATGTCAACCCCGCAAGATGGCCA 180
DB      308 GTGCGTGTCCAACCAAGGCGCAAGCACTGATGCAAGATGTCAACCCCGCAAGATGGCCA 367
QY      181 ACTTGGCCGGTAAACTGGGTTGGCTGACACCGCGGAGTGTGAGCTAATTCCTATCTC 240
DB      368 ACTTGGCCGGTAAACTGGGTTGGCTGACACCGCGGAGTGTGAGCTAATTCCTATCTC 427
QY      241 GGAACATTAACCAAGCGCAAGGTCTGTGAGCGCAAGATTCTGCGCAAAAAGGCTGCGCG 300
DB      428 GGAACATTAACCAAGCGCAAGGTCTGTGAGCGCAAGATTCTGCGCAAAAAGGCTGCGCG 487
QY      301 CAGTCTGTGTCAGGTGCGCATTAATGCTGATGTGACAGACGTAACCGGTAAACAGGCGCA 360
DB      488 CAGTCTGTGTCAGGTGCGCATTAATGCTGATGTGACAGACGTAACCGGTAAACAGGCGCA 547
QY      488 CAGTCTGTGTCAGGTGCGCATTAATGCTGATGTGACAGACGTAACCGGTAAACAGGCGCA 547
DB      548 GCTCTTCCGCGCGGCTGGCACTGACCTGTGCAATTAACGTAACCAAGGCAAGTACG 607
QY      421 TCCAGTCCCGATCCGTTATTTAATCTCTAATAAATCTGGGTTGGCAACTGATTAACGG 480
DB      608 TCCAGTCCCGATCCGTTATTTAATCTCTAATAAATCTGGGTTGGCAACTGATTAACGG 667
QY      481 AACGTGACTGAAGCATCTCTGACAGGCGAGGAGGTCAATTGCTGACTTTACCGGGCAT 540
```

```
DB      668 AACGTGACTGAAGCATCTCTGACAGGCGAGGAGGTCAATTGCTGACTTTACCGGGCAT 727
QY      541 CCGCAAAACGAGCGTTTGGCAACTGGAACGGGTGCTTAATTTTCCGAATCAAACTTGTGC 600
DB      728 CCGCAAAACGAGCGTTTGGCAACTGGAACGGGTGCTTAATTTTCCGAATCAAACTTGTGC 787
QY      601 CTTAAACGTGAATAACAGAACCAAGAACTTTCAATTAACGAGCATTTACCTGGAACTC 660
DB      788 CTTAAACGTGAATAACAGAACCAAGAACTTTCAATTAACGAGCATTTACCTGGAACTC 847
QY      661 AAGGTAGGCGCGCAATGCTCAATTAACCGGTGCGGTAGAGCTGCGCATCAATGTGACG 720
DB      848 AAGGTAGGCGCGCAATGCTCAATTAACCGGTGCGGTAGAGCTGCGCATCAATGTGACG 907
QY      721 GAGATATTTTCTCTGCAACAGCAAGGAATGCGAGCCGCGGTGGGGAAGATCAC 780
DB      908 GAGATATTTTCTCTGCAACAGCAAGGAATGCGAGCCGCGGTGGGGAAGATCAC 967
QY      781 GATTTCACACCATGTGAACACTTGTCTAATTTGTGCAATTAAGCGCAATTTTATTTGCTACAA 840
DB      968 GATTTCACACCATGTGAACACTTGTCTAATTTGTGCAATTAAGCGCAATTTTATTTGCTACAA 1027
QY      841 CGCAGCGCAGAGTGGCCCGCAGCGCGCACCCCGTTATTGGAATTGATTCAGGAGCG 900
DB      1028 CGCAGCGCAGAGTGGCCCGCAGCGCGCACCCCGTTATTGGAATTGATTCAGGAGCG 1087
QY      901 TTGACGCGCCCATCAACCGCAAAACAGGCGTATGTGTGACATTAACCACTTCAGTACTG 960
DB      1088 TTGACGCGCCCATCAACCGCAAAACAGGCGTATGTGTGACATTAACCACTTCAGTACTG 1147
QY      961 TTTATTTCCCGGACAGATATCTAATCTGTGCAATTTCTGGGCGGCACTGAGCTCACTG 1020
DB      1148 TTTATTTCCCGGACAGATATCTAATCTGTGCAATTTCTGGGCGGCACTGAGCTCACTG 1207
QY      1021 ACGCTTCCCGGTCAACCGCGGATTAACAGCGCGCAGGTGTGAACTGGTGTGAAACGCTGG 1080
DB      1208 ACGCTTCCCGGTCAACCGCGGATTAACAGCGCGCAGGTGTGAACTGGTGTGAAACGCTGG 1267
QY      1081 CGTGGCTTAAGCGATTAACGCGAGTGAATTCAGGTTTGTGCTGCTTCCAGACTTTACAG 1140
DB      1268 CGTGGCTTAAGCGATTAACGCGAGTGAATTCAGGTTTGTGCTGCTTCCAGACTTTACAG 1327
QY      1141 CAGATCGGTGAATAAACCGCGGTCTGATTAATTAAGCGCGCGGAGAGGTGAACCTGAC 1200
DB      1328 CAGATCGGTGAATAAACCGCGGTCTGATTAATTAAGCGCGCGGAGAGGTGAACCTGAC 1387
QY      1201 CTGGCAGATGTAAGAGGGAATGCGAGGCGCATGTGTGTTGTTGCAAGTGTTCAGCAA 1260
DB      1388 CTGGCAGATGTAAGAGGGAATGCGAGGCGCATGTGTGTTGTTGCAAGTGTTCAGCAA 1447
QY      1261 ATCGTAATGAAGCAGCATACCGGCGTGCAGTTTGAAT 1300
DB      1448 ATCGTAATGAAGCAGCATACCGGCGTGCAGTTTGAAT 1487

RESULT 8
US-09-866-379-9
/ Sequence 9, Application US/09866379
/ Patent No. 6855365
/ GENERAL INFORMATION:
/ APPLICANT: DIVERSA CORPORATION
/ APPLICANT: SHORT, Jay
/ APPLICANT: KREIZ, Keith
/ APPLICANT: GRAY, Kevin
/ APPLICANT: BARTON, Nelson
/ APPLICANT: GABRETT, James
/ APPLICANT: O'DONOHUE, Bileen
/ TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
/ FILE REFERENCE: DIVER1370-7
/ CURRENT APPLICATION NUMBER: US/09/866,379
/ PRIOR APPLICATION NUMBER: US 09/580,515
/ PRIOR FILING DATE: 2000-05-25
```

PRIOR APPLICATION NUMBER: US 09/318,528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: US 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 09/259,214
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: US 08/910,798
PRIOR FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.1
SEQ ID NO 9
LENGTH: 1901
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(1901)
OTHER INFORMATION: n is any nucleotide
US-09-866-379-9

Query Match 97.3%; Score 1286.2; DB 3; Length 1901;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1292; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

1 ATGAAAGGATCTTATCCATTTTATCTCTTGATTCGGTTAAACCCGCAATCTGCA 60
188 ATGAAAGGATCTTATCCATTTTATCTCTTGATTCGGTTAAACCCGCAATCTGCA 247
61 TTCGCTCAGAGTGAGCCGAGAGCTGAGAGTGAGAGTGAGTGTGATGTCAGTGTGATGAT 120
248 TTCGCTCAGAGTGAGCCGAGAGCTGAGAGTGAGAGTGAGTGTGATGTCAGTGTGATGAT 307
121 GTGCGTGTCCAAACCAAGCCCAAGCACTGATGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 180
308 GTGCGTGTCCAAACCAAGCCCAAGCACTGATGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 367
181 ACTGCGCGGTAAACTGGTGTGCTGACACCGCAGAGTGTGAGAGTGTGAGAGTGTGAGAG 240
368 ACTGCGCGGTAAACTGGTGTGCTGACACCGCAGAGTGTGAGAGTGTGAGAGTGTGAGAG 427
241 GAGCATTTACCAACGCGCAGCGTGTGAGCCGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 300
428 GAGCATTTACCAACGCGCAGCGTGTGAGCCGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 487
301 CAGTGTGTGATGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 360
488 CAGTGTGTGATGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 547
361 GCTTGTGCGCGCGGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 420
548 GCTTGTGCGCGCGGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 607
421 TCCAGTCCCGATCCGTTATTTAATCTCTTAAAACTGGCGTTTGCCTGATTAACGCG 480
608 TCCAGTCCCGATCCGTTATTTAATCTCTTAAAACTGGCGTTTGCCTGATTAACGCG 667
481 AACGCGATGACGCGATCTCTGACGAGGCGAGAGGCTCAATTGCTGATCTTAACCGGGCAT 540
668 AACGCGATGACGCGATCTCTGACGAGGCGAGAGGCTCAATTGCTGATCTTAACCGGGCAT 727
541 CGGCAAAAGCGGCTTTCGGAAGTGTGAGAGGCGGTTAATTTTCCGCAATCAAACTTGTGC 600
728 CGGCAAAAGCGGCTTTCGGAAGTGTGAGAGGCGGTTAATTTTCCGCAATCAAACTTGTGC 787
601 CTTAAACGTGAGAAACAGAGCAAAAGCTGTTCAATTACGACAGCATTTACATCGGAATC 660
788 CTTAAACGTGAGAAACAGAGCAAAAGCTGTTCAATTACGACAGCATTTACATCGGAATC 847
661 AAGGTGACCGCGCAAAATGTCTCAATTAACCGGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 720
848 AAGGTGACCGCGCAAAATGTCTCAATTAACCGGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 907
721 GAGATATTTCTCTGCAACAGCAAGGAGATGCGGAGCGGGGTGGGGAAGATCAC 780

908 GAGATATTTCTCTGCAACAGCAAGGAGATGCGGAGCGGGGTGGGGAAGATCAC 967
781 GATTCAACCAAGTGAACACCTTCTTAAGTTTGATTAACGCGCAATTTTATTTGCTACAA 840
968 GATTCAACCAAGTGAACACCTTCTTAAGTTTGATTAACGCGCAATTTTATTTGCTACAA 1027
841 CGCAGCGCAAGGTTTGGCCGAGCGCGCCACCCCGTTATTTGATTAATGATGACGACGCG 900
1028 CGCAGCGCAAGGTTTGGCCGAGCGCGCCACCCCGTTATTTGATTAATGATGACGACGCG 1087
901 TTGACGCCCATTCACCGCAAAACAGCGCTGATGATGATTAACCATCTTCACTGACTG 960
1088 TTGACGCCCATTCACCGCAAAACAGCGCTGATGATGATTAACCATCTTCACTGACTG 1147
961 TTATTTGCGGCAACAGTACTTAATCTGGAATATCTCGCGCGCACTGAGACTGACTG 1020
1148 TTATTTGCGGCAACAGTACTTAATCTGGAATATCTCGCGCGCACTGAGACTGACTG 1207
1021 ACGCTCCCGGTGACCGCGGATTAACAGCGCGCAGGTGTGAACTGTGTTGAACGCTGG 1080
1208 ACGCTCCCGGTGACCGCGGATTAACAGCGCGCAGGTGTGAACTGTGTTGAACGCTGG 1267
1081 CGTGGCTTAAGCATTAACAGCAGTGTGATTAAGTGTGCTGTCTTCCAGACTTTACAG 1140
1268 CGTGGCTTAAGCATTAACAGCAGTGTGATTAAGTGTGCTGTCTTCCAGACTTTACAG 1327
1141 CAGATGCGTGAATAACGCGCGCTGTCTTAATAATGCCCGCGGAGAGTGAACCTGACC 1200
1328 CAGATGCGTGAATAACGCGCGCTGTCTTAATAATGCCCGCGGAGAGTGAACCTGACC 1387
1201 CTGGCAGGATGTAAGAGCGGAATGCGCAGGAGCATGATGTTGTTGAGAGTTTACGCA 1260
1388 CTGGCAGGATGTAAGAGCGGAATGCGCAGGAGCATGATGTTGTTGAGAGTTTACGCA 1447
1261 ATCGTAATGAAGCAGCATACCGCGCTGACGTTTGAGAT 1300
1448 ATCGTAATGAAGCAGCATACCGCGCTGACGTTTGAGAT 1487

RESULT 9
US-09-866-379-6
Sequence 6, Application US/09866379
Patent No. 6855365
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: SHORT, Jay
APPLICANT: KREIZ, Keith
APPLICANT: GRAY, Kevin
APPLICANT: BARTON, Nelson
APPLICANT: GARRETT, James
APPLICANT: O'DONOHUE, Eileen
TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
FILE REFERENCE: DIVER1370-7
CURRENT APPLICATION NUMBER: US/09/866,379
CURRENT FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: US 09/580,515
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 09/318,528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: US 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 09/259,214
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: US 08/910,798
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 1901
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:

/ NAME/KEY: misc feature
/ LOCATION: (1)-(1901)
/ OTHER INFORMATION: n is any nucleotide
US-09-866-379-6

Query Match 97.0%; Score 1283; DB 3; Length 1901;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1290; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 ATGAAAGCATCTTAATCCCATTTTATCTCTTGTGATTCGGTTAACCCCGCAATCGCA 60
DB 188 ATGAAAGCATCTTAATCCCATTTTATCTCTTGTGATTCGGTTAACCCCGCAATCGCA 247

QY 61 TTGCGTCAGAGTGAAGCCGAGCTGAAGCTGAAGAGTGTGTATTCAGTGCCTCATGTGT 120
DB 248 TTGCGTCAGAGTGAAGCCGAGCTGAAGAGTGTGTATTCAGTGCCTCATGTGT 307

QY 121 GTGCGTGTCCAACCAAGGCGCACTGATGACAGATGTCAACCCGAGACGATGAGCA 180
DB 308 GTGCGTGTCCAACCAAGGCGCACTGATGACAGATGTCAACCCGAGACGATGAGCA 367

QY 181 ACCTGGCCGGTAAATCTGGGTTGGCTGACACCGCGNGGTGTGAGCTAATCGCTATCTC 240
DB 368 ACCTGGCCGGTAAATCTGGGTTGGCTGACACCGCGNGGTGTGAGCTAATCGCTATCTC 427

QY 241 GGACATTAACCAACGCGAGGCTGTGAGCGCAAGGATTCGTGCGAAAAAGGCGTCCCG 300
DB 428 GGACATTAACCTGGGCGCAAGCGCTGTGAGCGCAAGGATTCGTGCGAAAAAGGCGTCCCG 487

QY 301 CAGTCTGTGTCAGAGTCGCGATTAATGTCGTGATGTCAGACGATGATCCGTTAAACAGGCGAA 360
DB 488 CAGTCTGTGTCAGAGTCGCGATTAATGTCGTGATGTCAGACGATGATCCGTTAAACAGGCGAA 547

QY 361 GCGTTGCGCGCGGCGGCTGCGACCTGACTGTGCATTAACCGTACATACCCAGGCGATGAC 420
DB 548 GCGTTGCGCGCGGCGGCTGCGACCTGACTGTGCATTAACCGTACATACCCAGGCGATGAC 607

QY 421 TCCAGTCCCGGATCCGTTATTTAATCTCTTAAACCTGGGTTGCCAATGATTAACCGG 480
DB 608 TCCAGTCCCGGATCCGTTATTTAATCTCTTAAACCTGGGTTGCCAATGATTAACCGG 667

QY 481 AACGTGACTGACGCGATCTCTAGACAGGAGGAGGATCAATGCTGATTAACCGGCGAT 540
DB 668 AACGTGACTGACGCGATCTCTAGACAGGAGGAGGATCAATGCTGATTAACCGGCGAT 727

QY 541 CGGCAAAACGCGCTTTGCGCAACTGGAACGCGGTGCTTAATTTTCCGCAATCAACTTGTGC 600
DB 728 CGGCAAAACGCGCTTTGCGCAACTGGAACGCGGTGCTTAATTTTCCGCAATCAACTTGTGC 787

QY 601 CTTAAACGTGAAGAACAGAACGAAAGCTTTCAATTAACGACGATTAACATCGGAATC 660
DB 788 CTTAAACGTGAAGAACAGAACGAAAGCTTTCAATTAACGACGATTAACATCGGAATC 847

QY 661 AAGGTGAGGCGCGCAATGTCTCAATTAACGAGTGCCTGATGAGCTGATCAATGCTGAG 720
DB 848 AAGGTGAGGCGCGCAATGTCTCAATTAACGAGTGCCTGATGAGCTGATCAATGCTGAG 907

QY 721 GAGATATTTCTCTGCAACAGCAACAGGAATGCCGAGACCGGAGTGGGAAAGATCAC 780
DB 908 GAGATATTTCTCTGCAACAGCAACAGGAATGCCGAGACCGGAGTGGGAAAGATCAC 967

QY 781 GATTAACACATGGAACACCTTGTGATGATTAACGCGCAATTTTATTTGCTACAA 840
DB 968 GATTAACACATGGAACACCTTGTGATGATTAACGCGCAATTTTATTTGCTACAA 1027

QY 841 CGCAGCGCAGAGGTTCGCGCAGCGCGCGCAACCGGTTATTTGGAATTTGATGAGGAGCG 900
DB 1028 CGCAGCGCAGAGGTTCGCGCAGCGCGCGCGCAACCGGTTATTTGGAATTTGATGAGGAGCG 1087

QY 901 TTGACGCGCCCATCAACCGCAAAACAGGCGTATGCTGATCAATCAAGTCAAGTCA 960
DB 1088 TTGACGCGCCCATCAACCGCAAAACAGGCGTATGCTGATCAATCAAGTCAAGTCA 1147

QY 961 TTTATTCGCGGACAGATATCTAATCGGCAAAATCTGGGCGGCGCATGAGCTCAATCG 1020
DB 1148 TTTATTCGCGGACAGATATCTAATCGGCAAAATCTGGGCGGCGCATGAGCTCAATCG 1207

QY 1021 ACGCTTCCCGGTCAAGCGGATTAACAGCGCGCAGGTGTGAACCTGTGTTGAACGCTG 1080
DB 1208 ACGCTTCCCGGTCAAGCGGATTAACAGCGCGCAGGTGTGAACCTGTGTTGAACGCTG 1267

QY 1081 CGTGGCTTAAGGATTAACAGCGGATTAACAGCTTCCGCTGTCTTCAAGCTTTACAG 1140
DB 1268 CGTGGCTTAAGGATTAACAGCGGATTAACAGCTTCCGCTGTCTTCAAGCTTTACAG 1327

QY 1141 CAGATCGGATTAACAGCGGCTGTGATTAATTAAGCGCGCGGAGAGGATTAAGTAC 1200
DB 1328 CAGATCGGATTAACAGCGGCTGTGATTAATTAAGCGCGCGGAGAGGATTAAGTAC 1387

QY 1201 CTGGCAGATGTAAGAGGAAATGCGAGGCGATGTGCTGTGAGGATTTACGCA 1260
DB 1388 CTGGCAGATGTAAGAGGAAATGCGAGGCGATGTGCTGTGAGGATTTACGCA 1447

QY 1261 ATCGTAATGAAGCAGCATACCGGCGTCAAGTTGAGAT 1300
DB 1448 ATCGTAATGAAGCAGCATACCGGCGTCAAGTTGAGAT 1487

RESULT 10
US-09-540-149A-9
/ Sequence 9, Application US/09540149A
/ Patent No. 6511699
/ GENERAL INFORMATION:
/ APPLICANT: Lei, Xingren
/ TITLE OF INVENTION: ENZYMES WITH IMPROVED PHYTASE ACTIVITY
/ FILE REFERENCE: 19603/2791
/ CURRENT APPLICATION NUMBER: US/09/540,149A
/ CURRENT FILING DATE: 2000-03-31
/ PRIOR APPLICATION NUMBER: 60/127,032
/ PRIOR FILING DATE: 1999-03-31
/ NUMBER OF SEQ ID NOS: 9
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 9
/ LENGTH: 1489
/ TYPE: DNA
/ ORGANISM: Escherichia coli
US-09-540-149A-9

Query Match 95.6%; Score 1263.8; DB 3; Length 1489;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 1 ATGAAAGCATCTTAATCCCATTTTATCTCTTGTGATTCGGTTAACCCCGCAATCGCA 60
DB 182 ATGAAAGCATCTTAATCCCATTTTATCTCTTGTGATTCGGTTAACCCCGCAATCGCA 241

QY 61 TTGCGTCAGAGTGAAGCCGAGCTGAAGCTGAAGAGTGTGTATTCAGTGCCTCATGTGT 120
DB 242 TTGCGTCAGAGTGAAGCCGAGCTGAAGAGTGTGTATTCAGTGCCTCATGTGT 301

QY 121 GTGCGTGTCCAACCAAGGCGCACTGATGACAGATGTCAACCCGAGACGATGAGCA 180
DB 302 GTGCGTGTCCAACCAAGGCGCACTGATGACAGATGTCAACCCGAGACGATGAGCA 361

QY 181 ACCTGGCCGGTAAATCTGGGTTGGCTGACACCGGAGGTGTGAGCTAATCGCTATCTC 240
DB 362 ACCTGGCCGGTAAATCTGGGTTGGCTGACACCGGAGGTGTGAGCTAATCGCTATCTC 421

QY 241 GGACATTAACCAACGCGAGGCTGTGAGCGCAAGGATTCGTGCGAAAAAGGCGTCCCG 300
DB 422 GGACATTAACCAACGCGAGGCTGTGAGCGCAAGGATTCGTGCGAAAAAGGCGTCCCG 481

QY 301 CAGTCTGTGTCAGAGTCGCGATTAATGTCGTGATGTCAGACGATTAACCGTTAAACAGGCGAA 360
DB 482 CAGTCTGTGTCAGAGTCGCGATTAATGTCGTGATGTCAGACGATTAACCGTTAAACAGGCGAA 541

QY 841 CGCAGCCGAGAGTTCCCGCAGCCGCCACCCCGTTATTGGATTGGATTCATGGCAGCG 900
DB 1022 CGCAGCCGAGAGTTCCCGCAGCCGCCACCCCGTTATTGGATTGGATTCATGGCAGCG 1081
QY 901 TTGACGCCCCCATCCACCGCAAAAACAGCGGTATGTGTGACATTACCCACTTCAGTACTG 960
DB 1082 TTGACGCCCCCATCCACCGCAAAAACAGCGGTATGTGTGACATTACCCACTTCAGTACTG 1141
QY 961 TTTATTGCCGGACACGATATCTAATCTGGCAAAATCTCGCGCGCGCACTGAGCTCAACTG 1020
DB 1142 TTTATTGCCGGACACGATATCTAATCTGGCAAAATCTCGCGCGCGCACTGAGCTCAACTG 1201
QY 1021 AGCGTTCCCGGTTCAGCCCGGATTAACAGCCCGCGAGGTGTGAATGTGTGTTGAACCTG 1080
DB 1202 AGCGTTCCCGGTTCAGCCCGGATTAACAGCCCGCGAGGTGTGAATGTGTGTTGAACCTG 1261
QY 1081 CGTCGGCTAAGCGATTAACAGCCAGTGAATTCAGTTTCGTGTCTTCCAGACTTACAG 1140
DB 1262 CGTCGGCTAAGCGATTAACAGCCAGTGAATTCAGTTTCGTGTCTTCCAGACTTACAG 1321
QY 1141 CAGATCCGTGATTAACCGCGCTGTCAATTAATACCGCGCGCGAGAGGTGAACCTGACC 1200
DB 1322 CAGATCCGTGATTAACCGCGCTGTCAATTAATACCGCGCGCGAGAGGTGAACCTGACC 1381
QY 1201 CTGGCAGGATGTGAAGAGGAAATGCGCAGGGCATGTGTCTTGGCAGGTTTACGAA 1260
DB 1382 CTGGCAGGATGTGAAGAGGAAATGCGCAGGGCATGTGTCTTGGCAGGTTTACGAA 1441
QY 1261 ATCGTAATGAAGACGATACCGCGGTGCACTTTGAGATCTCATGAC 1308
DB 1442 ATCGTAATGAAGACGATACCGCGGTGCACTTTGAGATCTCATGAC 1489

RESULT 12
US-09-715-477-2
Sequence 2, Application US/09715477
Patent No. 6841370
GENERAL INFORMATION:
APPLICANT: Lei, Xingen
TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYLASE
FILE REFERENCE: 19603/4031
CURRENT APPLICATION NUMBER: US/09/715,477
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/166,179
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 1486
TYPE: DNA
ORGANISM: Escherichia coli
US-09-715-477-2

Query Match 95.5%; Score 1263; DB 3; Length 1486;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 1275; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 ATGAAGGATCTTATCCCATTTTATCTCTTGATTCGGTTAACCCCGCAATCTGCA 60
DB 188 ATGAAGGATCTTATCCCATTTTATCTCTTGATTCGGTTAACCCCGCAATCTGCA 247
QY 61 TTGCGTCAAGTGAAGCCGAGCTGAAAGCTGAAAGTGTGTGATTTGTCAGTGTCAATG 120
DB 248 TTGCGTCAAGTGAAGCCGAGCTGAAAGCTGAAAGTGTGTGATTTGTCAGTGTCAATG 307
QY 121 GTGCGGTCTCAACCAAGGCAAGCAATGATGCAAGATGCAACCCAGAGCATGGCCA 180
DB 308 GTGCGGTCTCAACCAAGGCAAGCAATGATGCAAGATGCAACCCAGAGCATGGCCA 367
QY 181 ACCTGGCCGGTAAACTGGGTGTGCTGACACCGCGAGGTGTGAGCTAATCGCTATCTC 240
DB 368 ACCTGGCCGGTAAACTGGGTGTGCTGACACCGCGAGGTGTGAGCTAATCGCTATCTC 427

QY 241 GGAATTAACCAAGCCAGGCTCTGGTAGCCGACGAAATGCTGCGCAAAAAGGCTGCCG 300
DB 428 GGAATTAACCAAGCCAGGCTCTGGTAGCCGACGAAATGCTGCGCAAAAAGGCTGCCG 487
QY 301 CAGTCGTGACAGGTGCGATTAATGCTGATGTGACAGAGCGTAACCGGTAAACAGCGCAA 360
DB 488 CAGTCGTGACAGGTGCGATTAATGCTGATGTGACAGAGCGTAACCGGTAAACAGCGCAA 547
QY 361 GCGTTGCGCGCGGCTGCGACCTGACCTGTGAATTAACCGTACATTAACCGAGCAGATAG 420
DB 548 GCGTTGCGCGCGGCTGCGACCTGACCTGTGAATTAACCGTACATTAACCGAGCAGATAG 607
QY 421 TCGAGTCCCGATCCGTTATTAATCTCTTAATAATCGCGTTTGCCCACTGATTAACGCG 480
DB 608 TCGAGTCCCGATCCGTTATTAATCTCTTAATAATCGCGTTTGCCCACTGATTAACGCG 667
QY 481 AACGTGACTGAAGCGATCTCTGACAGGGGAGGAGGTCAAATTGTCGACTTTACCGGCGAT 540
DB 668 AACGTGACTGAAGCGATCTCTGACAGGGGAGGAGGTCAAATTGTCGACTTTACCGGCGAT 727
QY 541 CGGCAAAAGCGCTTTGCGCAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTC 600
DB 728 CGGCAAAAGCGCTTTGCGCAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTC 787
QY 601 CTTAAACGTGAGAAAACAGACGAAAGCTGTCATTAACGACGACATTAACATCGAATCTC 660
DB 788 CTTAAACGTGAGAAAACAGACGAAAGCTGTCATTAACGACGACATTAACATCGAATCTC 847
QY 661 AAGGTAGAGCGCGCAATGCTCATTAACCGGTGAGGTAAAGCTGCGATCAATGCTGAAG 720
DB 848 AAGGTAGAGCGCGCAATGCTCATTAACCGGTGAGGTAAAGCTGCGATCAATGCTGAAG 907
QY 721 GAGATATTTCTCTGCAACAAGCAGGGAATGCGGAGCCGGGTGGGGAAGATCAAC 780
DB 908 GAGATATTTCTCTGCAACAAGCAGGGAATGCGGAGCCGGGTGGGGAAGATCAAC 967
QY 781 GATTCAACACAGTGAACAACCTTGCTAAGTTGCAATTAACGCGCAATTTTATTTGCTACA 840
DB 968 GATTCAACACAGTGAACAACCTTGCTAAGTTGCAATTAACGCGCAATTTTATTTACTACA 1027
QY 841 CGCAGCGCAGAGGTTCGCGCAGCGCGCACCCCGTTATGGATTGATATCAATGCAACG 900
DB 1028 CGCAGCGCAGAGGTTCGCGCAGCGCGCACCCCGTTATGGATTGATATCAATGCAACG 1087
QY 901 TTGACGCCCATCCACCGCAAAAACAGGCGTATGTGTGACATTAACCACTTCAGTGTG 960
DB 1088 TTGACGCCCATCCACCGCAAAAACAGGCGTATGTGTGACATTAACCACTTCAGTGTG 1147
QY 961 TTTATTGCCGGACACGATATCTAATCTGCGCAAAATCTCGCGCGGCACTGAGCTCAACTG 1020
DB 1148 TTTATTGCCGGACACGATATCTAATCTGCGCAAAATCTCGCGCGGCACTGAGCTCAACTG 1207
QY 1021 AGCGTTCCCGGTTCAGCCCGGATTAACAGCCGAGGTGTGAATGTGTGTTGAAGAGCTG 1080
DB 1208 AGCGTTCCAGGTTCAGCCCGGATTAACAGCCGAGGTGTGAATGTGTGTTGAAGAGCTG 1267
QY 1081 CGTCGGCTAAGCGATTAACAGCAAGTGAATTCAGGTTTGCTGTGCTTCCAGACTTTAAG 1140
DB 1268 CGTCGGCTAAGCGATTAACAGCAAGTGAATTCAGGTTTGCTGTGCTTCCAGACTTTAAG 1327
QY 1141 CAGATCCGTGATTAACCGCGCTGTCAATTAATACCGCGCGGAGAGGTGAACCTGACC 1200
DB 1328 CAGATCCGTGATTAACCGCGCTGTCAATTAATACCGCGCGGAGAGGTGAACCTGACC 1387
QY 1201 CTGGCAGGATGTGAAGAGGAAATGCGCAGGGCATGTGTGTTGGCAGGTTTACGCA 1260
DB 1388 CTGGCAGGATGTGAAGAGGAAATGCGCAGGGCATGTGTGTTGGCAGGTTTACGCA 1447
QY 1261 ATCGTAATGAAGACGATACCGCGGTGCACTTTG 1296
DB 1448 ATCGTAATGAAGACGATACCGCGGTGCACTTTG 1483

RESULT 13
US-09-715-477-4
; Sequence 4, Application US/09715477
; Patent No. 6841370
; GENERAL INFORMATION:
; APPLICANT: Lei, Xinggen
; TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYTASE
; FILE REFERENCE: 19603/4031
; CURRENT FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,179
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1486
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-715-477-4

Query Match 94.7%; Score 1251.8; DB 3; Length 1486;
Best Local Similarity 97.8%; Pred. No. 0;
Matches 1268; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 1 ATGAAGGAGTCTTAATCCATTTTATCTCTGATTCGTTAACCCCGCAATCTGCA 60
DB 188 ATGAAGGAGTCTTAATCCATTTTATCTCTGATTCGTTAACCCCGCAATCTGCA 247
QY 61 TTCGCTCAGAGTGAACCGGAGCTGAAAGCTGAAAGTGTGATTTGTCAGTCTCATGT 120
DB 248 TTCGCTCAGAGTGAACCGGAGCTGAAAGTGTGATTTGTCAGTCTCATGT 307
QY 121 GTGCGTCTCCCAACCAAGGCCCACTGATGCAAGATGTCACCCCGAGCATGCGCA 180
DB 308 GTGCGTCTCCCAACCAAGGCCCACTGATGCAAGATGTCACCCCGAGCATGCGCA 367
QY 181 ACCTGGCCGGTAAACTGGGTTGGCTGACACCGCGGAGTGGAGTAAATGCGCTATCTC 240
DB 368 ACCTGGCCGGTAAACTGGGTTGGCTGACACCGCGGAGTGGAGTAAATGCGCTATCTC 427
QY 241 GGACATTAACAACGCGCAGCTGTGCTGAGCCGAGATTGCTGCGCAAAAAGGCGTGC 300
DB 428 GGACATTAACAACGCGCAGCTGTGCTGAGCCGAGATTGCTGCGCAAAAAGGCGTGC 487
QY 301 CAGTCTGCTCAGGTGGCGATTATTCCTGATGTCAGACGCGTAAACCGTAAACAGGCGAA 360
DB 488 CAGTCTGCTCAGGTGGCGATTATTCCTGATGTCAGACGCGTAAACCGTAAACAGGCGAA 547
QY 361 GCTTGGCGCGCGGCTGGCACTGACCTGTGCAATTAACCGTAAACCGGCGATACG 420
DB 548 GCTTGGCGCGCGGCTGGCACTGACCTGTGCAATTAACCGTAAACCGGCGATACG 607
QY 421 TCCAGTCCCGATCCGTTATTTTAACTCTTAAAACTGGCGTTTGGCACTGGATTAACGCG 480
DB 608 TCCAGTCCCGATCCGTTATTTTAACTCTTAAAACTGGCGTTTGGCACTGGATTAACGCG 667
QY 481 AACGTGATGACCGCATCTCTCAGCAGGCGGAGAGGCTCAATTGCTGACTTTAACGGGCGAT 540
DB 668 AACGTGATGACCGCATCTCTCAGCAGGCGGAGAGGCTCAATTGCTGACTTTAACGGGCGAT 727
QY 541 CGGCAAAACGGGCTTTCGGGAAGTGGAGGGGCTTAATTTTCCGCAATCAAACTTGGAC 600
DB 728 CGGCAAAACGGGCTTTCGGGAAGTGGAGGGGCTTAATTTTCCGCAATCAAACTTGGAC 787
QY 601 CTAAACCTGAGAAACAGAGCAAAAGCTGTTCAATTAACGAGGCACTTAACATCGGAATC 660
DB 788 CTAAACCTGAGAAACAGAGCAAAAGCTGTTCAATTAACGAGGCACTTAACATCGGAATC 847
QY 661 AAGGTGAGCGCGCAATGCTCATTAACCGGTGCGTGAAGCTGCAATGCTGACG 720
DB 848 AAGGTGAGCGCGCAATGCTCATTAACCGGTGCGTGAAGCTGCAATGCTGACG 907

QY 721 GAGATATTTCTCCTGCAACAAGCAGAGGAATGCCGAGCGGGGTGGGAGATCAAC 780
DB 908 GAAATATTTCTCTGCAACAAGCAGAGGAATGCCGAGCGGGGTGGGAGATCAACT 967
QY 781 GATTACACCAAGTGAACACCTTGTCTAAGTTTGTGATTAACGCGCAATTTTATTTGCTCAA 840
DB 968 GATTACACCAAGTGAACACCTTGTCTAAGTTTGTGATTAACGCGCAATTTTATTTACTCAA 1027
QY 841 CGCACGCCAAGAGTTGCCCGGAGCGCGGCCACCCCGTTATTTGATTTGATCATGCGACG 900
DB 1028 CGCACGCCAAGAGTTGCCCGGAGCGCGGCCACCCCGTTATTTGATTTGATCATGCGACG 1087
QY 901 TTGACGCCCATCCACCGCAAAAACAGAGCTATGATGATTAACCACTTCAGTACTG 960
DB 1088 TTGACGCCCATCCACCGCAAAAACAGAGCTATGATGATTAACCACTTCAGTACTG 1147
QY 961 TTTATTCGCCGACACGATTAATCTGCAAAATCTGCGCGCGGACATGAGCTCAACTG 1020
DB 1148 TTTATTCGCCGACACGATTAATCTGCAAAATCTGCGCGCGGACATGAGCTCAACTG 1207
QY 1021 ACGCTTCCCGGTCAAGCCGCGATTAACAAGCCGCGCAGGTGTGTAATCTGTTGAACGCTGG 1080
DB 1208 ACGCTTCCCGGTCAAGCCGCGATTAACAAGCCGCGCAGGTGTGTAATCTGTTGAACGCTGG 1267
QY 1081 CGTGGCTAAGCGATTAACAGCAGTGAATTCAGGTTTCGCTGCTTCCAGACTTTACAG 1140
DB 1268 CGTGGCTAAGCGATTAACAGCAGTGAATTCAGGTTTCGCTGCTTCCAGACTTTACAG 1327
QY 1141 CAGATGCGTGAATAAACCGCGCTGTCAATTAATACGCGCGCGGAGAGGTGAACCTGACC 1200
DB 1328 CAGATGCGTGAATAAACCGCGCTGTCAATTAATACGCGCGCGGAGAGGTGAACCTGACC 1387
QY 1201 CTGGCAGATGTTGAAGCCGAAATGCGCAGGCGATGTGTTCTGTGCGAGGTTTAACGCA 1260
DB 1388 CTGGCAGATGTTGAAGCCGAAATGCGCAGGCGATGTGTTCTGTGCGAGGTTTAACGCA 1447
QY 1261 ATCGTAATGAAGCAGCATACCGCGCGTGCAGTTTG 1296
DB 1448 ATCGTAATGAAGCAGCATACCGCGCGTGCAGTTTG 1483

RESULT 14
US-08-910-798-1
; Sequence 1, Application US/08910798
; Patent No. 5876997
; GENERAL INFORMATION:
; APPLICANT: KRETZ
; TITLE OF INVENTION: NOVEL PHYTASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: August 13, 1997
; APPLICATION NUMBER: US/08/910,798
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HAILE, PH.D., LISA A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09010/029001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1272 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 IMMEDIATE SOURCE:
 CLONE: PHYTASB
 FEATURE:
 NAME/KEY:
 LOCATION:
 US-08-910-798-1

Query Match 91.5%; Score 1210; DB 2; Length 1272;
 Best Local Similarity 96.1%; Pred. No. 0;
 Matches 1212; Conservative 0; Mismatches 0; Indels 51; Gaps 1;

```

QY 1 ATGAAAGCGATCTTAATCCATTTTATCTCTGATTCCTTAACCCCGCAATCGCA 60
DB 1 ATGAAAGCGATCTTAATCCATTTTATCTCTGATTCCTTAACCCCGCAATCGCA 60
QY 61 TTCGCTCAGATGAGCCGGAAGCTGAAAGTGTGATGTTGTCAGTCTGATGGT 120
DB 61 TTCGCTCAGATGAGCCGGAAGCTGAAAGTGTGATGTTGTCAGTCTGATGGT 120
QY 121 GTGCGGTCTCCAAACCAAGCCAGCAATGATGACAGATGTCACCCCAAGCATGGCCA 180
DB 121 GTGCGGTCTCCAAACCAAGCCAGCAATGATGACAGATGTCACCCCAAGCATGGCCA 180
QY 181 ACCTGGCCGGTAAACCTGGGTTGGCTGACACCGCGNGGTGTGAGCTAAATCGCTATCTC 240
DB 181 ACCTGGCCGGTAAACCTGGGTTGGCTGACACCGCGNGGTGTGAGCTAAATCGCTATCTC 240
QY 241 GGAATTTACCAACGCCAGCGCTGTGATGCGGACGAAATGCTGGCGAAAAAGGCTGCCG 300
DB 241 GGAATTTACCAACGCCAGCGCTGTGATGCGGACGAAATGCTGGCGAAAAAGGCTGCCG 300
QY 301 CAGTCTGTGACAGTCCGCAATTAATGCTGATGTCAGACAGCGTACCCGTAAACAGGCGAA 360
DB 301 CAGTCTGTGACAGTCCGCAATTAATGCTGATGTCAGACAGCGTACCCGTAAACAGGCGAA 360
QY 361 GCGTTCCGCGCGGCTGGACCTGACTGTGCAATAACGTACATACCAGGACGATAGC 420
DB 361 GCGTTCCGCGCGGCTGGACCTGACTGTGCAATAACGTACATACCAGGACGATAGC 420
QY 421 TCCAGTCCCGATCCGTTATTTAATCTCTAATAAACTGGGCTTGGCACTGATTAACGGC 480
DB 421 TCCAGTCCCGATCCGTTATTTAATCTCTAATAAACTGGGCTTGGCACTGATTAACGGC 480
QY 481 AACGTGACTGACGCAATCTCTGACGACGAGGAGGCTCAATTGCTGACTTTAACCGGCACT 540
DB 481 AACGTGACTGACGCAATCTCTGACGACGAGGAGGCTCAATTGCTGACTTTAACCGGCACT 540
QY 541 CGGCAAAACGGCTTTCGCGCAACTGAAACGGGTGCTTAATTTTCCGCAATCAACTGTGTC 600
DB 541 CGGCAAAACGGCTTTCGCGCAACTGAAACGGGTGCTTAATTTTCCGCAATCAACTGTGTC 600
QY 601 CTTAAACGTGAAACAGGACGAAAGCTGTTCATTAACGACAGCATTAACATCGAACTC 660
DB 601 CTTAAACGTGAAACAGGACGAAAGCTGTTCATTAACGACAGCATTAACATCGAACTC 660
QY 661 AAGGTGAGGCGCGAATGCTCATTAACCGGTGCGGTAAAGCTCGCATCATATGCTGAGC 720
DB 661 AAGGTGAGGCGCGAATGCTCATTAACCGGTGCGGTAAAGCTCGCATCATATGCTGAGC 720
QY 721 GAGATATTTCTCTGCAACAAGCAACAGGAATGCGGAGCCGAGGTGGGAAAGATCAC 780
DB 721 GAGATATTTCTCTGCAACAAGCAACAGGAATGCGGAGCCGAGGTGGGAAAGATCAC 780
QY 781 GATTACACCAAGTGAACACTTGTCTAAGTTTGCATTAACGCGCAATTTTATTTGCTACAA 840
DB 781 GATTACACCAAGTGAACACTTGTCTAAGTTTGCATTAACGCGCAATTTTATTTGCTACAA 840
QY 841 CGACGCCACAGAGTTGCTCCGCAAGCCGCGCACCCCGTTATTGGATTGATCATGACGCG 900
DB 841 CGACGCCACAGAGTTGCTCCGCAAGCCGCGCACCCCGTTATTGGATTGATCATGACGCG 900
QY 901 TTGAGCCCCCATTCACCGCAAAAACAGGAGTATGAGTGAACATTACCCACTTCAGTAC 960
DB 901 TTGAGCCCCCATTCACCGCAAAAACAGGAGTATGAGTGAACATTACCCACTTCAGTAC 960
QY 961 TTTATTTCCGGAACACATATCTAATCTGGCAAACTTCGCGCGGCACTGAGACTCACTG 1020
DB 961 TTTATTTCCGGAACACATATCTAATCTGGCAAACTTCGCGCGGCACTGAGACTCACTG 1020
QY 1021 ACGCTTCCCGGTTCAGCCGATTAACAGCCCGCAAGGTGTGAACCTGTTTGAACCTG 1080
DB 1021 ACGCTTCCCGGTTCAGCCGATTAACAGCCCGCAAGGTGTGAACCTGTTTGAACCTG 1080
QY 970 ACGCTTCCCGGTTCAGCCGATTAACAGCCCGCAAGGTGTGAACCTGTTTGAACCTG 1029
DB 970 ACGCTTCCCGGTTCAGCCGATTAACAGCCCGCAAGGTGTGAACCTGTTTGAACCTG 1029
QY 1081 CGTGGCTAAGCGATTAACAGCCAGTGAATTCAGGTTTGTGCTGTCTTCAGACTTACAG 1140
DB 1081 CGTGGCTAAGCGATTAACAGCCAGTGAATTCAGGTTTGTGCTGTCTTCAGACTTACAG 1140
QY 1141 CAGATGCGTGAATTAACCGCGCTGTCTAATTAATACGCGCCGCGAGAGGTGAACCTGAC 1200
DB 1141 CAGATGCGTGAATTAACCGCGCTGTCTAATTAATACGCGCCGCGAGAGGTGAACCTGAC 1200
QY 1201 CTGGCAGATGTGAAGAGGAAATGCGACGGGCAATGTGTTGTTGGCAGGTTTACGCA 1260
DB 1201 CTGGCAGATGTGAAGAGGAAATGCGACGGGCAATGTGTTGTTGGCAGGTTTACGCA 1260
QY 1150 CTGGCAGATGTGAAGAGGAAATGCGACGGGCAATGTGTTGTTGGCAGGTTTACGCA 1209
DB 1150 CTGGCAGATGTGAAGAGGAAATGCGACGGGCAATGTGTTGTTGGCAGGTTTACGCA 1209
QY 1261 ATCGTAATGAAGACGCGTACCGGCGTGTGAGTTCATCAACCATCAATCAC 1320
DB 1261 ATCGTAATGAAGACGCGTACCGGCGTGTGAGTTCATCAACCATCAATCAC 1320
QY 1210 ATCGTAATGAAGACGCGTACCGGCGTGTGAGTTCATCAACCATCAATCAC 1269
DB 1210 ATCGTAATGAAGACGCGTACCGGCGTGTGAGTTCATCAACCATCAATCAC 1269
QY 1321 TAA 1323
DB 1270 TAA 1272

```

RESULT 15

US-09-489-039A-341
 / Sequence 341, Application US/09489039A
 / Patent No. 6610836
 / GENERAL INFORMATION:
 / APPLICANT: Gary Breton et. al
 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 / TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 / FILE REFERENCE: 2709.2004001
 / CURRENT APPLICATION NUMBER: US/09/489,039A
 / CURRENT FILING DATE: 2000-01-27
 / PRIOR APPLICATION NUMBER: US 60/117,747
 / PRIOR FILING DATE: 1999-01-29
 / NUMBER OF SEQ ID NOS: 14342
 / SEQ ID NO 341
 / LENGTH: 1266
 / TYPE: DNA
 / ORGANISM: Klebsiella pneumoniae
 / US-09-489-039A-341

Query Match 4.8%; Score 63.6; DB 3; Length 1266;
 Best Local Similarity 48.7%; Pred. No. 1.7e-10;
 Matches 202; Conservative 0; Mismatches 210; Indels 3; Gaps 1;

```

QY 86 AGCTGAAAAGTGTGATTTGTCAGTGTGATGCTGCTGCTCC--AACCAAGGCCA 142
DB 113 AGCTGACACAGGTGTGATTAAGCCCTCAACCTGCGTGGCCGCTTGGCCATAACG 172
QY 143 CGCAATGATGAGGATGTCACCCAGACGATGCGCAACCTGCGCGGTAAACCTGGGTT 202
DB 173 GCAAGTCTGTGACATGTCACCGCAAGGCTTGGCCGCAAGTGTGCGGCGGCGC 232
QY 203 GAGTGAACCGCGAGGTGTGAGCTAATGCTTATCTCGACATTACCAAGCCAGCGTC 262
DB 223 AGCTGACACCAAGGCGGTGTGCTGAGAGTGTATATGGGCACTAATATGGCGAATGGC 292

```


Publication No. US20060099612A1
GENERAL INFORMATION:
APPLICANT: SUNTORY LIMITED
APPLICANT: NAKAO, YOSHIHIRO
APPLICANT: NAKAMURA, NORIHISA
APPLICANT: KODAMA, YUKIKO
APPLICANT: FUJIMURA, TOMOKO
APPLICANT: ASHIKARI, TOSHIHIKO
TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
FILE REFERENCE: S-38-285
CURRENT APPLICATION NUMBER: US/11/217,529
CURRENT FILING DATE: 2005-09-02
PRIOR APPLICATION NUMBER: US 10/932,182
PRIOR FILING DATE: 2004-09-02
NUMBER OF SEQ ID NOS: 197023
SOFTWARE: PatentIn version 3.3
SEQ ID NO 79920
LENGTH: 1257
TYPE: DNA
ORGANISM: Saccharomyces pastorianus
US-11-217-529-79920

Query Match 2.7%; Score 34.8; DB 7; Length 1257;
Best Local Similarity 65.4%; Pred. No. 0.061; 27; Indels 0; Gaps 0;
Matches 51; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 671 CCGACTGTGTCTCATTAACCGGTGCGGTGAGCTTCGATCAATGCTGACGAGATATTTTC 730
DB 253 CCGAGTCTGTCTCAAAAATGGTGTGTGAGCGTCCATCAAGCGCTGAAGAGCCGTATTT 194
QY 731 TCCTGCAACAGACAGG 748
DB 193 TCATCAATCCGTAGAG 176

RESULT 3
US-11-293-697-2180/C
Sequence 2180, Application US/11293697
Publication No. US20060105376A1
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: Novel full length cDNA
FILE REFERENCE: H1-A0106
CURRENT APPLICATION NUMBER: US/11/293,697
CURRENT FILING DATE: 2005-12-05
PRIOR APPLICATION NUMBER: US/10/108,260
PRIOR FILING DATE: 2002-03-28
NUMBER OF SEQ ID NOS: 5458
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2180
LENGTH: 2879
TYPE: DNA
ORGANISM: Homo sapiens
US-11-293-697-2180

Query Match 2.6%; Score 34; DB 7; Length 2879;
Best Local Similarity 50.6%; Pred. No. 0.19; 80; Indels 0; Gaps 0;
Matches 82; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 134 CCAAGCCACCGCACTGATGACAGATGTCACCCAGACGATGCGCAACTGCGCGTAA 193
DB 388 CCGCGCCCTTGTATTGAGAGAGAGGTGTCCCAAGGCGTCTCGGCGCCGCAAGCT 329
QY 194 AACTGGGTGAGTGAACCGCGCGGTGTGAGCTATGCTATCTCGACATTACTGCGC 253
DB 328 GCATGTGTGGGGAAGTCCGCCAGCGGGGCGCTGATCATCTCGGCGGTGAAGTCCGCTC 269
QY 254 GTCAGCGTGTGATGACCGACGAGTGTGCTCAATATGTGCT 255
DB 268 GGAACGGCGCTTGAGTGACCGAGCTGACACCAAGTTGCT 227

RESULT 4

US-10-953-349-11004/C
Sequence 11004, Application US/10953349
Publication No. US20060107345A1
GENERAL INFORMATION:
APPLICANT: ALEXANDROV, Nickolai et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
FILE REFERENCE: 2750-1579PUS2
CURRENT APPLICATION NUMBER: US/10/953,349
CURRENT FILING DATE: 2004-09-30
NUMBER OF SEQ ID NOS: 40252
SOFTWARE: PatentIn version 3.3
SEQ ID NO 11004
LENGTH: 2271
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-10-953-349-11004

Query Match 2.6%; Score 33.4; DB 6; Length 2271;
Best Local Similarity 60.4%; Pred. No. 0.27; 36; Indels 0; Gaps 0;
Matches 55; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 1212 TGAAGAGGAAATGCGAGGCGATGTGCTGTCGACAGGTTTACGCAATGTGAATGA 1271
DB 1734 TGAACGGGACATGTTCGGGTGATATGATCACTGAAGAGGTTTCTGCAATCTGAAGA 1675
QY 1272 AGCAGCATACCGGCGTGCATTTGAGATCT 1302
DB 1674 CCGGAAACACTAGTTGCTCTTGATCTTT 1644

RESULT 5
US-10-953-349-27515/C
Sequence 27515, Application US/10953349
Publication No. US20060107345A1
GENERAL INFORMATION:
APPLICANT: ALEXANDROV, Nickolai et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
FILE REFERENCE: 2750-1579PUS2
CURRENT APPLICATION NUMBER: US/10/953,349
CURRENT FILING DATE: 2004-09-30
NUMBER OF SEQ ID NOS: 40252
SOFTWARE: PatentIn version 3.3
SEQ ID NO 27515
LENGTH: 1393
TYPE: DNA
ORGANISM: Trifolium aestivum
US-10-953-349-27515

Query Match 2.5%; Score 32.4; DB 6; Length 1393;
Best Local Similarity 48.4%; Pred. No. 0.46; 96; Indels 0; Gaps 0;
Matches 90; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 603 TAAAGTGAGAAAGAGAGCAAGAGCTGTTCAATTAACGAGGATTAACCATCGAACTCAA 662
DB 1167 TAACTTTAGTACATCTTGAGGTTTCAACCAAGTAAGTAAAGCTGACAGTCCA 1108
QY 663 GGTGAGCCCGACTGTGTCTCATTAACCGGTGCGGTAAAGCTTCGATCAATGTGACGGA 722
DB 1107 GGGGTGGGGCGGTGGGGGTGAGTGAACAAGCTCAAGTCACTATTCAGTGTCCCGG 1048
QY 723 GATATTTTCTGCAACAGACAGGAAATGCGGAGCGGGGTGGGAAGATACCGGA 782
DB 1047 GATTAAGTCACTAACATACACAGGTCCGTTGAAGACAGAAAGGAGTCTCACGA 988
QY 783 TTCACA 788
DB 987 TTTCACA 982

RESULT 6
US-10-449-902-8987

```
/ Sequence 8987, Application US/10449902
/ Publication No. US20060123505A1
/ GENERAL INFORMATION:
/ APPLICANT: National Institute of Agricultural Sciences.
/ APPLICANT: Bio-oriented Technology Research Advancement Institution.
/ APPLICANT: The Institute of Physical and Chemical Research.
/ APPLICANT: Foundation for Advancement of International Science.
/ TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
/ FILE REFERENCE: MOA-A02051-US
/ CURRENT APPLICATION NUMBER: US/10/449,902
/ PRIOR FILING DATE: 2003-05-29
/ PRIOR APPLICATION NUMBER: JP 2002-203269
/ PRIOR FILING DATE: 2002-05-30
/ PRIOR APPLICATION NUMBER: JP 2002-383870
/ PRIOR FILING DATE: 2002-12-11
/ NUMBER OF SEQ ID NOS: 56791
/ SOFTWARE: Patent In Ver. 2.1
/ SEQ ID NO 8987
/ LENGTH: 1396
/ TYPE: DNA
/ ORGANISM: Oryza sativa
/ PUBLICATION INFORMATION:
/ DATABASE ACCESSION NUMBER: AK106335
/ DATABASE ENTRY DATE: 2002-08-28
US-10-449-902-8987

Query Match      2.4%; Score 32; DB 6; Length 1396;
Best Local Similarity 58.3%; Pred. No. 0.63;
Matches 56; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 679 GTCCTATTACCGGCGGTGTAAGCCCTCGCATGATGCTGACGAGATATTCTCTCGCAA 738
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 456 GTCTGTTCTCGGTGCGGTGCTCCCTCTTCTGCGGAGCGGAGAGCTGCTACTCCGAGAA 515
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 739 CAAGCAGACGAGATGCGCGAGCGCGGGGTGGGAGAG 774
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 516 CCGGCATGGGAGACAGCGCGCGCGAGCATGTGTGCGG 551
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7
US-10-449-902-13868
/ Sequence 13868, Application US/10449902
/ Publication No. US20060123505A1
/ GENERAL INFORMATION:
/ APPLICANT: National Institute of Agricultural Sciences.
/ APPLICANT: Bio-oriented Technology Research Advancement Institution.
/ APPLICANT: The Institute of Physical and Chemical Research.
/ APPLICANT: Foundation for Advancement of International Science.
/ TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
/ FILE REFERENCE: MOA-A02051-US
/ CURRENT APPLICATION NUMBER: US/10/449,902
/ PRIOR FILING DATE: 2003-05-29
/ PRIOR APPLICATION NUMBER: JP 2002-203269
/ PRIOR FILING DATE: 2002-05-30
/ PRIOR APPLICATION NUMBER: JP 2002-383870
/ PRIOR FILING DATE: 2002-12-11
/ NUMBER OF SEQ ID NOS: 56791
/ SOFTWARE: Patent In Ver. 2.1
/ SEQ ID NO 13868
/ LENGTH: 2634
/ TYPE: DNA
/ ORGANISM: Oryza sativa
/ PUBLICATION INFORMATION:
/ DATABASE ACCESSION NUMBER: AK11216
/ DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-13868

Query Match      2.4%; Score 31.8; DB 6; Length 2634;
Best Local Similarity 46.9%; Pred. No. 1.1;
Matches 99; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 199 GGTGAGCTGACACCGGCGGTGTGAGCAATACGCTATCTGAGCATTAATGAGGTGAG 258
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
DB 317 GGCCTTGCGGCTCCACGTCGCGGTCTCGAGAGGCGCGCTCGACGTCGCGCCGCTCTG 376
QY 259 GGTCTGTAACCGGACGAGATTGCTGCTTAATATGTGCTGCCGCACTGTGTAGGTCCG 318
DB 377 CTTCTGTGAGGCGAGCGGTGTATCTCGGCGGGTTCCGCCCGCACCTCGACGCCGGGA 436
QY 319 ATTATGTGTAATGTCGAGAGCGTACCCGCTAAACAGCGCGAAAGCTTGGCGCGGGCTG 378
DB 437 CTTCTGCTCGTCAACCGCATCTCTCTGTGAGGGGAGCAGAGGTGTTCAAGCCGAGCAA 496
QY 379 GCACCTGACTGTGCAATACCGTACATACCC 409
DB 497 CGAGCTGATCTGCAAGAGCAGCGCATGAC 527

RESULT 8
US-10-449-902-13863/c
/ Sequence 13863, Application US/10449902
/ Publication No. US20060123505A1
/ GENERAL INFORMATION:
/ APPLICANT: National Institute of Agricultural Sciences.
/ APPLICANT: Bio-oriented Technology Research Advancement Institution.
/ APPLICANT: The Institute of Physical and Chemical Research.
/ APPLICANT: Foundation for Advancement of International Science.
/ TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
/ FILE REFERENCE: MOA-A02051-US
/ CURRENT APPLICATION NUMBER: US/10/449,902
/ PRIOR FILING DATE: 2003-05-29
/ PRIOR APPLICATION NUMBER: JP 2002-203269
/ PRIOR FILING DATE: 2002-05-30
/ PRIOR APPLICATION NUMBER: JP 2002-383870
/ PRIOR FILING DATE: 2002-12-11
/ NUMBER OF SEQ ID NOS: 56791
/ SOFTWARE: Patent In Ver. 2.1
/ SEQ ID NO 13863
/ LENGTH: 3129
/ TYPE: DNA
/ ORGANISM: Oryza sativa
/ PUBLICATION INFORMATION:
/ DATABASE ACCESSION NUMBER: AK11211
/ DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-13863

Query Match      2.4%; Score 31.4; DB 6; Length 3129;
Best Local Similarity 47.7%; Pred. No. 1.7;
Matches 92; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 999 CGCGCACTGAGAGCTCAACTGAGAGCTTCCCGGTCAAGCGGATTAACGCCGCCAGGTG 1058
DB 2942 CGAAGTTCGAGAGCTCAACTGCGCTGCGCATGTGCTTCCGTCGCGCCGTTGAATGGGG 2883
QY 1059 TGAACGTGTGTTTGAACGCTGCGCTGAGGATTAACGCAAGTGTGATTCAGGTTTC 1118
DB 2882 TCGCGTGAACCGGTTACCGAATTTGAAGCCAGAGGTGATGCGGCTGTGGAAGCCATC 2823
QY 1119 GCTGCTTTCAGACTTTACAGCAGATGCGTATTAACGCCGCTGTCTATTAAATACGCC 1178
DB 2822 GGTGAACCGGAGATGTTGAGAGCATGAGAGAGAACGCCGCGGCATTTGCGGAGGGA 2763
QY 1179 GCCCGAGAGGTTG 1191
DB 2762 ACCAGAGTGTCTG 2750

RESULT 9
US-10-449-902-10760/c
/ Sequence 10760, Application US/10449902
/ Publication No. US20060123505A1
/ GENERAL INFORMATION:
/ APPLICANT: National Institute of Agricultural Sciences.
/ APPLICANT: Bio-oriented Technology Research Advancement Institution.
/ APPLICANT: The Institute of Physical and Chemical Research.
/ APPLICANT: Foundation for Advancement of International Science.
```

;; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
;; FILE REFERENCE: MOA-A0205Y1-US
;; CURRENT APPLICATION NUMBER: US/10/449,902
;; CURRENT FILING DATE: 2003-05-29
;; PRIOR APPLICATION NUMBER: JP 2002-203269
;; PRIOR FILING DATE: 2002-05-30
;; PRIOR APPLICATION NUMBER: JP 2002-383870
;; PRIOR FILING DATE: 2002-12-11
;; NUMBER OF SEQ ID NOS: 56791
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 10760
;; LENGTH: 1344
;; TYPE: DNA
;; ORGANISM: Oryza sativa
;; PUBLICATION INFORMATION:
;; DATABASE ACCESSION NUMBER: AK108108
;; DATABASE ENTRY DATE: 2002-08-28
US-10-449-902-10760

Query Match 2.4%; Score 31; DB 6; Length 1344;
Best Local Similarity 53.8%; Pred. No. 1.4;
Matches 64; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 258 GCGTCTGCTAGCCGACGAGTGTGCTTAAATGTGCTGCCGACGATCTGTGTCAGTCCG 317
|||
Db 289 GCGTCTGCTAGCTCGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 230
Qy 318 GATTATGCTGATGTCGACGACGCTTACCCGTAACAGCCGACGCTTCCGCGCGGCG 376
|||
Db 229 TGTGTTGCTGCTGTCGACGAGGAGTCCGCAACACAGAGAGCGCGCTCCGCGGAC 171

RESULT 10
US-11-293-697-707/c
; Sequence 707, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 707
; LENGTH: 2791
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-707

Query Match 2.4%; Score 31; DB 7; Length 2791;
Best Local Similarity 55.0%; Pred. No. 2.2;
Matches 61; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 59 CATTCGCTCAGAGTACGCGGAGCTGAGCGTGAAGGTGTGATTTCTCAGTGTATG 118
|||
Db 1419 CTTTCGACGTAGTGTATACGCTCTTAAAGCGGTGTGTGAGAGTTGTTCAATCTCTCTG 1360
Qy 119 GTGTCGCTGCTCAACGACGACGCACTGATGTCAGAGTCAACCCAG 169
|||
Db 1359 GTGGTTCGTGTCTCTGCTGCTCAGAGTGAAGCTCAGATCTTCCGG 1309

RESULT 11
US-10-449-902-15013
; Sequence 15013, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.

;; APPLICANT: Foundation for Advancement of International Science.
;; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
;; FILE REFERENCE: MOA-A0205Y1-US
;; CURRENT APPLICATION NUMBER: US/10/449,902
;; CURRENT FILING DATE: 2003-05-29
;; PRIOR APPLICATION NUMBER: JP 2002-203269
;; PRIOR FILING DATE: 2002-05-30
;; PRIOR APPLICATION NUMBER: JP 2002-383870
;; PRIOR FILING DATE: 2002-12-11
;; NUMBER OF SEQ ID NOS: 56791
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 15013
;; LENGTH: 3553
;; TYPE: DNA
;; ORGANISM: Oryza sativa
;; PUBLICATION INFORMATION:
;; DATABASE ACCESSION NUMBER: AK065548
;; DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-15013

Query Match 2.4%; Score 30.8; DB 6; Length 3553;
Best Local Similarity 47.8%; Pred. No. 3;
Matches 89; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

Qy 297 CCGCAGTCTGTCTCAGGTCGCGATTAATGCTGATGTCGACGAGCTACCCGTAACAGG 356
|||
Db 2688 CTCGATTTTGACTTGCTGCAAGCTTATGATGAAAAGAACCATATGACGAAAT 2747
Qy 357 CGAAGCTTTCGCGCGCGGCTGCGACCTGACCTGCAATACCGTACATACCCAGGAGA 416
|||
Db 2748 CGCTGCGACATTCGCGCTATCTGCGACCCGAGTATGCAATGAGAGGCCATCTGACTGAAA 2807
Qy 417 TAGTCGACATCCGATCCGATTAATTAATCTCTTAAACAGGCTTCCCACTGATAA 476
|||
Db 2808 GCGTATGTTTTCGATTTGAGATTGTTGCTCTAGAAACAGTTGCGCGATCAACAC 2867
Qy 477 CGCGAA 482
|||
Db 2868 CGACAA 2873

RESULT 12
US-10-517-441-65/c
; Sequence 65, Application US/10517441
; Publication No. US20060121467A1
; GENERAL INFORMATION:
; APPLICANT: FOERKENS, John
; APPLICANT: HARBECK, Nadia
; APPLICANT: KOENIG, Thomas
; APPLICANT: MATER, Sabine
; APPLICANT: MARTENS, John
; APPLICANT: MODEL, Fabian
; APPLICANT: NIMMICH, Inko
; APPLICANT: RUJAN, Tamas
; APPLICANT: SCHMITT, Armin
; APPLICANT: SCHMITT, Manfred
; APPLICANT: LOOK, Maxime P.
; APPLICANT: MARX, Almuth
; APPLICANT: HOEFER, Heinz
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell]
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; CURRENT FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: DE 10317955.0
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: DE 10300096.8
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: DE 10245779.4
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 2147

SEQ ID NO 65
LENGTH: 3050
TYPE: DNA
ORGANISM: Homo Sapiens
US-10-517-441-65

Query Match
Best Local Similarity 2.3%; Score 30.6; DB 6; Length 3050;
Matches 60; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 160 GTACCCGAGACGATGCGCACTGCGGTTAAACCTGAGCTGACACCGCGCGT 219
DB 281 GCCGCTCCCGTCCGATCATCCGCGGCTGAGCGGCCCGCGCAGGC 222
QY 220 GGTGAGCTAATGCGCTATTCGACATTACTGCGCGTCTGTAG 268
DB 221 TGTGGCTAACCTCCATCTTAGATCCACCAGCGGAGGAG 173

RESULT 13
US-10-449-902-965/c
Sequence 965, Application US/10449902
Publication No. US20060123505A1

GENERAL INFORMATION:
APPLICANT: National Institute of Agricultural Sciences.
APPLICANT: Bio-oriented Technology Research Advancement Institution.
APPLICANT: The Institute of Physical and Chemical Research.
TITLE OF INVENTION: Foundation for Advancement of International Science.
FILE REFERENCE: MOA-A0205Y1-US
CURRENT APPLICATION NUMBER: US/10/449,902
CURRENT FILING DATE: 2003-05-29
PRIOR APPLICATION NUMBER: JP 2002-203269
PRIOR FILING DATE: 2002-05-30
PRIOR APPLICATION NUMBER: JP 2002-383870
PRIOR FILING DATE: 2002-12-11
NUMBER OF SEQ ID NOS: 56791
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 965
LENGTH: 422
TYPE: DNA
ORGANISM: Oryza sativa
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: AK059042
DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-965

Query Match
Best Local Similarity 2.3%; Score 30.2; DB 6; Length 422;
Matches 65; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 658 CTCAGGTGAGCGCGAGCTGTCTCACTTAACCGGTGCGGTAAAGCTGCATCAATGCTG 717
DB 224 CTCAGGTGAGCTGTGTATTCGCGCACCATGATGAAGAAACACGACCAAAAATG 165
QY 718 ACGGAGATTTCTCTGCAACAGCAAGGAAATGCCGAGCCGGGTGGGAAAGATC 777
DB 164 ATGAGGAAGATCTTAGGCGCGGAGACAGCGCGGAGCTCCCTGACGCGAGCGGAGAG 105
QY 778 ACC 780
DB 104 ACC 102

RESULT 14
US-10-449-902-18030
Sequence 18030, Application US/10449902
Publication No. US20060123505A1
GENERAL INFORMATION:
APPLICANT: National Institute of Agricultural Sciences.
APPLICANT: Bio-oriented Technology Research Advancement Institution.
APPLICANT: The Institute of Physical and Chemical Research.
APPLICANT: Foundation for Advancement of International Science.

TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
FILE REFERENCE: MOA-A0205Y1-US
CURRENT APPLICATION NUMBER: US/10/449,902
CURRENT FILING DATE: 2003-05-29
PRIOR APPLICATION NUMBER: JP 2002-203269
PRIOR FILING DATE: 2002-05-30
PRIOR APPLICATION NUMBER: JP 2002-383870
PRIOR FILING DATE: 2002-12-11
NUMBER OF SEQ ID NOS: 56791
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 18030
LENGTH: 772
TYPE: DNA
ORGANISM: Oryza sativa
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: AK068443
DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-18030

Query Match
Best Local Similarity 2.3%; Score 30.2; DB 6; Length 772;
Matches 110; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 137 AGGCCAGCACTGATGAGATGTCACCCAGAGCATGCGCAACTGGCCGTTAAAC 196
DB 50 AGGCCCTCTCTCTGTGACCTGTTCAACGACAGCGGCGCGCGGAGGCG 109
QY 197 TGGTGAAGCTACACCGCGGTGAGTAACTGCTATCTCGGACATTAATGCGCTC 256
DB 110 CGGGCGGTGACGATGCGGCGGAGAGACGCGTGGCGGAGCTGGCCGACGTCGCGC 169
QY 257 AGCGTGTGAGCCGACGATGCTGCTTAATGTGCGCCGACGTCGTGTCAGTGC 316
DB 170 GCGGTGCGGCGCGGCGGAGATGCGCGTCCGACGAAATCATCTGTCAGCAGATCG 229
QY 317 CGATTATGCTGATGTCACGAGCATGACCGTAAACAGCGGAACTTCGCCCGCGGC 376
DB 230 CGCGCGGTGAGAGAGCGCGCGTCCGCGCGCGCGCGCGCGCGCGCGCGC 289
QY 377 TGG 379
DB 290 AGG 292

RESULT 15
US-11-289-102-50/c
Sequence 50, Application US/11289102
Publication No. US2006012151A1
GENERAL INFORMATION:

APPLICANT: Lee, Hyerin
APPLICANT: Shaw, Peter M.
APPLICANT: Clark, Edwin
TITLE OF INVENTION: BIOMARKERS AND METHODS FOR DETERMINING SENSITIVITY TO
TITLE OF INVENTION: MICROTUBULE-STABILIZING AGENTS
FILE REFERENCE: 10338 NP
CURRENT APPLICATION NUMBER: US/11/289,102
CURRENT FILING DATE: 2005-11-29
PRIOR APPLICATION NUMBER: US 60/631,993
PRIOR FILING DATE: 2004-11-30
NUMBER OF SEQ ID NOS: 395
SOFTWARE: PatentIn version 3.3
SEQ ID NO 50
LENGTH: 3452
TYPE: DNA
ORGANISM: Homo sapiens
US-11-289-102-50

Query Match
Best Local Similarity 2.3%; Score 30; DB 7; Length 3452;
Matches 45; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 997 GCGGCGCATGAGCTCACTGACCGCTTCCCGTACGCCGATACAGCCCGCAGCT 1056

Db	417	GACGTCGCTTGAGCCCCGTCCTAGACCTCGCTGGCTGAGGGTGATGACATGCCGCAAGT	358
Qy	1057	GGTGAAGTGG	1066
Db	357	GGGGCACTGG	348

Search completed: June 14, 2006, 14:23:14
Job time : 39.7719 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2006, 10:01:41 ; Search time 6667.62 Seconds
(without alignments)
10650.337 Million cell updates/sec

Title: US-10-601-319-9
Perfect score: 1308
Sequence: 1 atgaagcgatcatcattcc.....gcagttcgatccatccta 1308

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 2795965780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: gb_est1:*
2: gb_est3:*
3: gb_est4:*
4: gb_est5:*
5: gb_est6:*
6: gb_hic:*
7: gb_est2:*
8: gb_est7:*
9: gb_est8:*
10: gb_est9:*
11: gb_g861:*
12: gb_g862:*
13: gb_g863:*
14: gb_g864:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query length	DB ID	Description
1	592	45.3	616	2	BU074127 BU074127
2	580.2	44.4	853	13	CU662734 PR10142b
3	531.4	40.6	746	4	CA093060 SCCCL200
4	524.2	40.1	529	3	BU113770 SJAAUBG02
5	488.2	37.3	595	5	CF326092 JMT1--05-
6	483.2	36.9	895	8	CN762997 ID0AA5DC
7	446.8	34.2	868	8	CN754382 ID0AA12D
8	442.8	33.9	864	8	CN757565 ID0AA12D
9	372.4	28.5	393	7	BE520240 M1B125TM
10	370.2	28.3	383	7	BE520240 M1B125TM
11	370.2	28.3	383	7	AM036132 EST74510
12	348.6	26.7	531	3	AM036134 EST74510
13	295.8	22.6	706	8	BU113769 SJAAUBG01
14	295.4	22.6	354	1	CN759004 ID0AA24B
15	254.4	19.4	707	2	AA545747 HBKSP184-
16	198	15.1	320	7	BU619443 BU619443
17	193.6	14.8	208	7	BE436403 EST407481
18	193.6	14.8	208	2	BM409016 EST58343
19	189.2	14.5	546	2	BM412806 EST587133
					BU040796 BU040796

20	160.8	12.3	657	2	BG457132	BG457132 NF100C05P
21	159.8	12.2	179	7	BE461872	BE461872 EST413291
22	157.4	12.0	1171	8	CV672454	CV672454 RET75J_01
23	151.2	11.6	508	2	BU030203	BU030203 BU030203
24	143.2	10.9	254	7	BE520241	BE520241 M1B12XTM
25	100.2	7.7	769	8	CO386885	CO386885 AGENCOURT
26	88.8	6.8	863	8	CV672453	CV672453 RET75J_01
27	59.8	4.6	772	14	DX010757	DX010757 KB+BDJCO
28	59.4	4.5	821	13	CU665831	CU665831 PRI1044E
29	59.4	4.5	842	13	CU665745	CU665745 PRI1044E
30	54	4.1	470	9	DN201839	DN201839 USDA-FP_1
31	52.6	4.0	1175	10	CK943192	CK943192 LamjGeetG
32	43.8	3.3	392	9	HE0613	HE0613 YF44d12_81
33	41.8	3.2	274	3	BO153360	BO153360 NF03C071
34	40.8	3.1	250	14	DX065699	DX065699 KBEB073U0
35	39.6	3.0	909	4	CA117609	CA117609 SCBFLR108
36	39.4	3.0	600	4	CB547796	CB547796 AMGNNUC:C
37	39.2	3.0	553	7	AM390601	AM390601 MR3-ST019
38	39	3.0	733	2	BI152280	BI152280 602917706
39	38.8	3.0	1160	9	DR148985	DR148985 49350064
40	37.8	2.9	939	12	CU078996	CU078996 CH216-154
41	37.6	2.9	347	4	CB705089	CB705089 AMGNNUC:M
42	37.6	2.9	402	3	BU671248	BU671248 NISC_1r06
43	37.6	2.9	418	5	CK739626	CK739626 USDA-FP_6
44	37.6	2.9	425	4	CB696209	CB696209 AMGNNUC:N
45	37.6	2.9	499	3	BP490202	BP490202 BP490202

ALIGNMENTS

RESULT 1
BU074127/c 616 bp mRNA linear EST 29-SEP-2003
LOCUS BU074127 NIBB Mochii normalized Xenopus tailbud library Xenopus
DEFINITION laevis cDNA clone XJ090j06 5', mRNA sequence.
ACCESSION BU074127.1 GI:17504316
VERSION BU074127.1
KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Xenopus.
1 (bases 1 to 616)
Kikayama,A., Texasaka,C., Mochii,M., Ueno,N., Shin-I,T. and Kohara,Y.
Expressed genes in X. laevis embryo
Unpublished (2001)
Contact: Tadasu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@gene.nig.ac.jp
The information of this clone is available through the following URL.
http://xenopus.nibb.ac.jp.
Location/Qualifiers
1..616
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="XJ090j06"
/issue_type="whole embryo"
/dev_stage="stage 25"
/clone_id="NIBB Mochii normalized Xenopus tailbud library"

ORIGIN
Query Match 45.3%; Score 592; DB 2; Length 616;
Best Local Similarity 97.6%; Pred. No. 4.9e-175;

```
Matches 601; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 264 GGTGACCGACGATTTGCTGCTTAATGTGGCTGCCCGGAGTCTGTGTCAGGTCGGGATTTAT 323
Db 616 GGTGACCGACGATTTGCTGCGGAAAAAGGGCTGCCCGGAGTCTGTGTCAGGTCGGGATTTAT 557
QY 324 TGTGATGTGACGACGAGGTACCCCGTAAAAACAGGCGAAGCCCTTCGCGCGCGGGCTGGACAC 383
Db 556 TGTGATGTGACGACGAGGTACCCCGTAAAAACAGGCGAAGCCCTTCGCGCGCGGGCTGGACAC 497
QY 384 TGACTGTGACATTAACCGTACATACCAGCAGATACGTCCAGTCCCGATCCGTTATTTAA 443
Db 496 TGACTGTGACATTAACCGTACATACCAGCAGATACGTCCAGTCCCGATCCGTTATTTAA 437
QY 444 TCTCTTAAAACTGGCGTTTGGCACTGAGATTAAGCGCAAGGTGACTGACGCGATCTTGA 503
Db 436 TCTCTTAAAACTGGCGTTTGGCACTGAGATTAAGCGCAAGGTGACTGACGCGATCTTGA 377
QY 504 GAGGCGAGGAGGTCAATTGCTGACTTAACGGGCGATTATGAAAACGGCGTTTCGGCAACT 563
Db 376 CAGGCGAGGAGGTCAATTGCTGACTTAACGGGCGATTATGAAAACGGCGTTTCGGCAACT 317
QY 564 GAAACGGGTGCTTAATTTTCCGCAATCAACTTGTGCTTAAACGTGAAAAACAGAACGA 623
Db 316 GAAACGGGTGCTTAATTTTCCGCAATCAACTTGTGCTTAAACGTGAAAAACAGAACGA 257
QY 624 AAGCTGTTCATTAACGACGAGATTAACATTCGGAATCTGAAGTGAAGCGCGCATGTGTCTC 683
Db 256 AAGCTGTTCATTAACGACGAGATTAACATTCGGAATCTGAAGTGAAGCGCGCATGTGTCTC 197
QY 684 ATTAACCGGTGCGGTAAAGCTTCGATCAATGCTGACGAGATTAATTTCTCTGCAACAAGC 743
Db 196 ATTAACCGGTGCGGTAAAGCTTCGATCAATGCTGACGAGATTAATTTCTCTGCAACAAGC 137
QY 744 ACAGGGAATGCGGAGCGGAGTGGGAAAGATCACCGGATTCACACGAGTGAACACCTT 803
Db 136 ACTGGGAATGCGGAGCGGAGTGGGAAAGATCACCGGATTCACACGAGTGAACACCTT 77
QY 804 GCTAAGTTGATTAACGCGCAATTTGATTGCTACACGACGCGGAGGTTGCGCGGAG 863
Db 76 GCTAAGTTGATTAACGCGCAATTTGATTGCTACACGACGCGGAGGTTGCGCGGAG 17
QY 864 CCGGCGCAACCCCGTTA 879
Db 16 CCGGCGCAACCCCGTTA 1

RESULT 2
CL662734 853 bp DNA linear GSS 09-JUL-2004
LOCUS PR10142b.B11 - PR10142b.B21 (853) Mixed stage fosmid library of P.
DEFINITION pacificus var. California Pristionchus pacificus genomic, genomic
survey sequence.
ACCESSION CL662734
VERSION CL662734.1 GI:50150877
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Pristionchus pacificus
Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
REFERENCE 1 (bases 1 to 853)
AUTHORS Strinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
TITLE Apprads: an Acedb database for the nematode satellite organism
JOURNAL Pristionchus pacificus
PUBMED Nucleic Acids Res. 32 (1), D421-D422 (2004)
COMMENT 14681447
Contact: Sommer RJ
Evolutinary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel.: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@uebingen.mpg.de
```

```
This library was generated at Caltech, Pasadena, USA and end
sequenced at Vancouver, Canada.
Seq primer: T7
Class: fosmid ends.
Location/Qualifiers
1. 853
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:94126"
/clone_1fb="Mixed stage fosmid library of P. pacificus
var. California"
/notes="Vector: pBf105-5 Fosmid vector"

ORIGIN
Query Match 44.4%; Score 580.2; DB 13; Length 853;
Best Local Similarity 98.7%; Pred. No. 2.8e-11;
Matches 585; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 708 ATCAATGCTGACGAGATTAATTTCTCTGCAACAAGCAGGGAATGCCGAGCGGGGTG 767
Db 1 ATCAATGCTGACGAGATTAATTTCTCTGCAACAAGCAGGGAATGCCGAGCGGGGTG 60
QY 768 GGGAGGATCAACCGATTCAACACAGTGAACACCTTGTCTAAGTTGCATTAACGCAATT 827
Db 61 GGGAGGATCAACCGATTCAACACAGTGAACACCTTGTCTAAGTTGCATTAACGCGCAATT 120
QY 828 TGATTTGCTACACGACGCGCAGAGTGGCCGAGCGCGGACCCCGGTTATTAAGTTT 887
Db 121 TTAATTTCTACACGACGCGCAGAGTGGCCGAGCGCGGACCCCGGTTATTAAGTTT 180
QY 888 GATCAAGACAGCTTTGACGCGCCATCCACCGCAAAAACAGGCGATGATGATGACATTACC 947
Db 181 GATCAAGACAGCTTTGACGCGCCATCCACCGCAAAAACAGGCGATGATGATGACATTACC 240
QY 948 CACTTACGTGCTGTTTATCGCGGACAGATTAATCTGCAAAATCTGCGCGGCGCACT 1007
Db 241 CACTTACGTGCTGTTTATCGCGGACAGATTAATCTGCAAAATCTGCGCGGCGCACT 300
QY 1008 GGAAGCTCAACTGACGCGCTTCGCGGTCAAGCGGATTAACGCGCGGAGTGTGAACCTG 1067
Db 301 GGAAGCTCAACTGACGCGCTTCGCGGTCAAGCGGATTAACGCGCGGAGTGTGAACCTG 360
QY 1068 GTTTGAAGCGTGGCGGTGAGTGAAGCATTAACGCGGATTAATGAGTTTGGTGTGCTT 1127
Db 361 GTTTGAAGCGTGGCGGTGAGTGAAGCATTAACGCGGATTAATGAGTTTGGTGTGCTT 420
QY 1128 CCAACTTTTACAGCAGATGCGGTGAATAAAGCGCGTGTCTAATTAATAGCGCGCGGAGA 1187
Db 421 CCAACTTTTACAGCAGATGCGGTGAATAAAGCGCGTGTCTAATTAATAGCGCGCGGAGA 480
QY 1188 GGTGAACCTGACCTTGGCAGATGTGAAGAGCAAAATGCCGAGGCGATGTGTTGGC 1247
Db 481 GGTGAACCTGACCTTGGCAGATGTGAAGAGCAAAATGCCGAGGCGATGTGTTGGC 540
QY 1248 AGTTTAAAGCAATTCGTAATGAAGACGATACCGGCGGTGACGTTTGAAT 1300
Db 541 AGTTTAAAGCAATTCGTAATGAAGACGATACCGGCGGTGACGTTTGAAT 593

RESULT 3
LOCUS CA093060 746 bp mRNA linear EST 23-SEP-2003
DEFINITION SCCCL2001D10.b CL2 Saccharum officinarum cDNA clone SCCCL2001D10
3', mRNA sequence.
ACCESSION CA093060
VERSION CA093060.1 GI:34946367
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Saccharum officinarum
Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
```

REFERENCE 1 (bases 1 to 746)
 Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
 The libraries that made SUCEST
 Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
 CONTACT: Arruda P
 Centro de Biologia Molecular e Engenharia Genética
 Universidade Estadual de Campinas
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil
 Tel: 55 19 3788 1137
 Fax: 55 19 3788 1089
 Email: pattinda@unicamp.br
 Clone distribution: clone distribution information can be found
 through the Brazilian Clone Collection Center (BCCC) at
 http://www.bcccenter.fcav.unesp.br
 Plate: 001 row: D column: 10
 Seq primer: M13/Forward primer.
 Location/Qualifiers

FEATURES
 source 1..746
 /organism="Saccharum officinarum"
 /mol_type="mRNA"
 /db_xref="taxon:4547"
 /clone="SCCCL2001D10"
 /lab_host="XLIBLue MRP"
 /clone_1lb="CL2"
 /note="Organ: Pool of sugarcane calli submitted to low
 (40C) and high (37 C) temperature stress; Vector:
 pBluscript; Site_1: EcoRI; Site_2: XhoI; An
 undirectional cDNA library generated from [pool of
 sugarcane calli submitted to low (40C) and high (37 C)
 temperature stress]. cDNA was prepared from polyA+ mRNA
 using ZAP - cDNA Synthesis Kit (Stratagene). The
 double-strand cDNAs were fractionated in a sepharose CL-2B
 40cm-column and fragments sizing between 0.8 and 1.5 Kb
 were directionally cloned into the vector. Details of
 each source of RNA and library construction can be
 obtained at http://succest.lad.ic.unicamp.br/public"

ORIGIN

Query Match 40.6%; Score 531.4; DB 4; Length 746;
 Best Local Similarity 97.3%; Pred. No. 7.3e-156;
 Matches 549; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

OR 555 TTGTCCTTAAACGTGAGAAACAGACGAACTGTCTTAAACGAGCATTCATCATG 654
 DB 184 TTGCGTACTGTGAGAGAAACAGACGAACTGTCTTAAACGAGCATTCATCATG 243
 OR 655 GAACTCAAGGTGAGCGCCGACTGTGTCTTAAACGAGCATTCATCATG 714
 DB 244 GAACTCAAGGTGAGCGCCGACTGTGTCTTAAACGAGCATTCATCATG 303
 OR 715 CTGACGAGATATTTCTCTGCAACAGACGAGATGCGGAGCGGGGTGGGAAAG 774
 DB 304 CTGACGAGATATTTCTCTGCAACAGACGAGATGCGGAGCGGGGTGGGAAAG 363
 OR 775 ATCACCAGTTCAACAGGTGAAACCTTGTCTTAAACGAGCATTCATCATG 834
 DB 364 ATCACCAGTTCAACAGGTGAAACCTTGTCTTAAACGAGCATTCATCATG 423
 OR 835 CTACAGCAGCGCAGAGGTGCGCGAGCGCGCAACCGGTTATTGATTGATCAAG 894
 DB 424 CTACAGCAGCGCAGAGGTGCGCGAGCGCGCAACCGGTTATTGATTGATCAAG 483
 OR 895 ACAGGTTGACGCGCCATTCACCGCAAAAGCGGTATGTGTGACATTACCACTTCA 954
 DB 484 ACAGGTTGACGCGCCATTCACCGCAAAAGCGGTATGTGTGACATTACCACTTCA 543
 OR 955 GTGCTGTTATGCGCGGACAGATCTTATCTGGCAATCTGGGCGGCGCATCGAAGTC 1014
 DB 544 GTGCTGTTATGCGCGGACAGATCTTATCTGGCAATCTGGGCGGCGCATCGAAGTC 603
 OR 1015 AACTGAGCGCTCCCGGTGAGCGGATTAACGCGCGCGAGTGAATGCTGTGTTGA 1074

DB 604 AACTGAGCGCTCCCGGTGAGCGGATTAACGCGCGCGAGTGAATGCTGTGTTGA 662
 OR 1075 CGCTGCGGTGCGCTTAAGGATTAACGCGCGAGTTCAGTTTCTGCTGTTCCAGACT 1134
 DB 663 CGCTGCGGTGCGCTTAAGGATTAACGCGCGAGTTCAGTTTCTGCTGTTCCAGACT 722
 OR 1135 TTACAGCAGATGCGTATTAACG 1158
 DB 723 TTACAGCAGATGCGTATTAACG 746

RESULT 4
 BU713770
 LOCUS
 DEFINITION
 BU713770 529 bp mRNA linear EST 23-OCT-2003
 SUBMITTER Adult SJC 7/94 Schistosoma japonicum cDNA similar to
 pDblIDEX/A chain A, Crystal Structure Of Phylate Complex Of
 Escherichia Coli Phylate At Ph 6.6. Phylate Is Bound With Its
 3-Phosphate In The Active Site: Hg2+ Cation Acts As An
 Intermolecular Bridge, mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 Hu, W., Yan, Q., Shen, D.K., Liu, F., Zhu, Z. D., Song, H. D., Xu, X. R.,
 Wang, Z. J., Rong, Y. P., Zeng, L. C., Wu, J., Zhang, X., Wang, J. J.,
 Xu, X. N., Wang, S. Y., Fu, G., Zhang, X. L., Wang, Z. O., Birdley, P. J.,
 McManus, D. P., Xue, C. L., Feng, Z., Chen, Z. and Han, Z. G.
 Evolutionary and biomedical implications of a Schistosoma japonicum
 complementary DNA resource
 Nat. Genet. 35 (2), 139-147 (2003)

JOURNAL
 PUBMED
 COMMENT
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 201203, P. R. China
 Tel: 86-21-50801919 (ex. 45)
 Fax: 86-21-50801922
 Email: hanzg@chgc.sh.cn.

FEATURES
 source 1..529
 Location/Qualifiers

/organism="Schistosoma japonicum"
 /mol_type="mRNA"
 /strain="Chinese (Anhui) strain"
 /db_xref="taxon:6182"
 /sex="Male and female"
 /tissue_type="Whole body"
 /dev_stage="Adult worms"
 /lab_host="Mouse and rabbit"
 /clone_1lb="Adult SJC 7/94"
 /note="Vector: lambda ZAP-II XR; Site_1: EcoR I; Site_2:
 XhoI I; Several hundred adult Schistosoma japonicum
 (Anhui, P. R. China, strain), of mixed sex, were perfused
 from the mesenteries of experimentally infected mice and
 rabbits at the Queensland Institute of Medical Research,
 Brisbane, Australia (QIMR), and stored for several months
 in liquid nitrogen. Subsequently, mRNA was isolated at the
 QIMR from livers of these worms by oligo dT
 chromatography, using a kit from Pharmacia. The mRNA was
 then shipped to Clontech, Palo Alto, CA, USA, who
 constructed a cDNA library. First strand synthesis was
 primed with an oligo-dT-XhoI-primer and synthesized using
 M-MuV reverse transcriptase. Second strand synthesis was
 accomplished with RNase H and T4 DNA polymerase. The
 double stranded cDNA was ligated to EcoRI linkers,
 digested with EcoRI and XhoI, and ligated into the
 phagemid vector lambda ZAP II XR. After construction of
 this directional library by Clontech, it was returned to
 the QIMR. During analysis of the library at the QIMR, we
 have found that a small percentage, 28 to 3%, of the

clones contain inserts that appear to be highly homologous to sequences from salmonid fishes, as determined by homology comparisons using BLAST and by Southern hybridization analysis to genomic DNA from salmon (Sigma Chemical Co., St. Louis, MO) under stringent washing conditions. The remainder of the clones appear to contain *S. japonicum* sequences."

ORIGIN

Query Match 40.1%; Score 524.2; DB 3; Length 529;
Best Local Similarity 99.4%; Pred. No. 1.3e-15;
Matches 526; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

670 GCCGACCTGTCTCTAATTAACGGGTCGGGTAAAGCTTCAGATCAATGCTGACGGAATATT 729
Db 1 GCCGACAAATGTCTCTAATTAACGGGTCGGGTAAAGCTTCAGATCAATGCTGACGGAATATT 60
Qy 730 CTCCTGCAACAAGCAACAGGAATGCGGAGCGGGGTGGGAAGGATCACCATTCAAC 789
Db 61 CTCCTGCAACAAGCAACAGGAATGCGGAGCGGGGTGGGAAGGATCACCATTCAAC 120
Qy 790 CAGTGAACACCTTGTCTAAGTTTGCAATACCGCAATTTGATTGCTACAAAGCAGCCCA 849
Db 121 CAGTGAACACCTTGTCTAAGTTTGCAATACCGCAATTTGATTGCTACAAAGCAGCCCA 180
Qy 850 GAGGTGCGCGGAGCGCGGACCCGTTATTAGATTGATCAAGACGCGTTGACGCGC 909
Db 181 GAGGTGCGCGGAGCGCGGACCCGTTATTAGATTGATCAAGACGCGTTGACGCGC 240
Qy 910 CATCAACCGCAAAAACAGCGTATGATGATGATTAACCACTTCACTGCTGTTATTCGCC 969
Db 241 CATCAACCGCAAAAACAGCGTATGATGATGATTAACCACTTCACTGCTGTTATTCGCC 300
Qy 970 GGAACCACTTAACTTCTGCAATTCGCGCGCGGACCTGAGCTCAATCTGAACGTTCCC 1029
Db 301 GGAACCACTTAACTTCTGCAATTCGCGCGCGGACCTGAGCTCAATCTGAACGTTCCC 360
Qy 1030 GGTGACCGGATTAACAGCGCGGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 1089
Db 361 GGTGACCGGATTAACAGCGCGGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 420
Qy 1090 AGCGATTAACAGCGAGTGAATCAAGTTTCGCTGCTTTCAGACTTTACAGCAATGCGT 1149
Db 421 AGCGATTAACAGCGAGTGAATCAAGTTTCGCTGCTTTCAGACTTTACAGCAATGCGT 480
Qy 1150 GATTAACGCGGCTGTCTTAAATACGCGCGCGGAGAGGTGAATCTGA 1198
Db 481 GATTAACGCGGCTGTCTTAAATACGCGCGCGGAGAGGTGAATCTGA 529

RESULT 5
CP326092/c 595 bp mRNA linear EST 18-AUG-2003
LOCUS JMT1--05-B09.g1 AtJMT-overexpressing transgenic rice lambda phage
DEFINITION cDNA library (JMT1) Oryza sativa (japonica cultivar-group) cDNA
clone JMT1--05-B09, mRNA sequence.
ACCESSION CP326092.1 GI:33800445
VERSION CP326092.1 GI:33800445
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Bep
Clade; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE
AUTHORS Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
TITLE Large-scale Sequencing Analysis of Rice ESTs
JOURNAL Unpublished (2003)
COMMENT Contact: Nahm B.H.
Genomics and Genetics Institute, GreenGene Biotech Inc., Division
of Bioscience and Bioinformatics, Myongji University
Yongin, Kyeonggi, Korea

Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.
Location/Qualifiers

FEATURES

Source

1. 595
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/culivar="Nackdong"
/db_xref="taxon:39947"
/clone="JMT1--05-B09"
/tissue_type="leaf"
/dev_stage="14 days after germination"
/lab_host="E. coli SOLR"
/clone_lib="AtJMT-overexpressing transgenic rice lambda
phage cDNA library (JMT1)"
/note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
XhoI; cDNA was inserted into lambda Uni-ZAP XR vector at 5'
end with EcoRI and 3' end with XhoI site. mRNA was
prepared from Arabidopsis jasmonate Carboxyl
methyltransferase overexpression line."

ORIGIN

Query Match 37.3%; Score 488.2; DB 5; Length 595;
Best Local Similarity 99.4%; Pred. No. 3.1e-142;
Matches 490; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

808 AGTTTGATTAACGGCAATTTGATTTGCTCAACAGCAGCGAGGTTCGCCGAGCGC 867
Db 595 AGTTTGATTAACGGCAATTTGATTTGCTCAACAGCAGCGAGGTTCGCCGAGCGC 536
Qy 868 GCCACCCCGTTATAGATTGATCAAGACAGCGTTGACGCCCATCCACGCAAAAACAG 927
Db 535 GCCACCCCGTTATAGATTGATCAAGACAGCGTTGACGCCCATCCACGCAAAAACAG 476
Qy 928 GCGTATGATGATCAATCACTTCACTGCTGTTATGCGCGGACAGATTAATCTG 987
Db 475 GCGTATGATGATCAATCACTTCACTGCTGTTATGCGCGGACAGATTAATCTG 416
Qy 988 GCAATTCGCGCGCGGACGAGCTCACTGACGCTTCGCCGTCAGCGGATTAACG 1047
Db 415 GCAATTCGCGCGCGGACGAGCTCACTGACGCTTCGCCGTCAGCGGATTAACG 356
Qy 1048 CCGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1107
Db 355 CCGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 296
Qy 1108 ATTGAGGTTTCGCTGCTTTCAGACTTTACAGCAATGCGGTAAACGCGCTGTCA 1167
Db 295 ATTGAGGTTTCGCTGCTTTCAGACTTTACAGCAATGCGGTAAACGCGCTGTCA 236
Qy 1168 TTAAATACGCGCGCGGAGAGGTGAATCTGACCTGCGAGAGTGAAGAGCAAAATGCG 1227
Db 235 TTAAATACGCGCGCGGAGAGGTGAATCTGACCTGCGAGAGTGAAGAGCAAAATGCG 176
Qy 1228 CAGGCGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1287
Db 175 CAGGCGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 116
Qy 1288 TGCAATTGAGAT 1300
Db 115 TGCAATTGATAT 103

RESULT 6
CN762997/c 895 bp mRNA linear EST 20-MAY-2004
LOCUS ID0AA5DC09BML ApMS Acyrthosiphon pisum cDNA clone ID0AA5DC09 5',
DEFINITION mRNA sequence.
ACCESSION CN762997.1 GI:47536920
VERSION CN762997.1 GI:47536920
KEYWORDS EST.
SOURCE Acyrthosiphon pisum (pea aphid)
ORGANISM Acyrthosiphon pisum

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes; Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.

1 (bases 1 to 895)
Hunter, W., Martinez-Torres, D., Rabhe, Y., Sabater-Munoz, B., Stern, D., Tagu, D. and Wincker, P.
An expressed sequence tags database for the pea aphid *Acyrthosiphon pisum*

Unpublished (2004)
Contact: D. Tagu

JOURNAL
INRA Rennes
UMR BIO3P, BP 35327, F-35653 Le Rheu Cedex France
Tel: +33.2.23.48.51.65
Fax: +33.2.23.48.51.50
Risk of contamination by bacterial sequences from obligatory (Buchnera) or facultative endosymbionts. These sequences were obtained in the frame of the International Consortium of Aphid Genomics in collaboration with Genoscope

PCR Primers
FORWARD: CAGGAACGCTATGACC
Plate: 5

row: C column: 9.
Location/Qualifiers
1. 895
/organism="Acyrthosiphon pisum"
/mol_type="mRNA"
/cultivar="developmentstage"
/db_xref="taxon:7029"
/clone="ID0AA12DE01"
/tissue_type="whole insect"
/dev_stage="nymphs and adults (parthenogenetic females)"
/lab_host="XLI-Blue"
/clone_1lb="ApMS"
/note="Vector: pBS-SK minus, Site 1: EcoRI; Site 2: XhoI; Sample name: ID0AA1; Plant growth place: Department of Ecology & Evolutionary Biology, Princeton University; Soil conditions: Soil; Sowing date: 01/06/1999; Harvesting date: 01/06/1999; Stress date: no stress; Description: Aphids inoculated on one-week old *Vicia faba* under non-sterile conditions. All parthenogenetic stages and both winged and wingless adults were collected for library construction. ; experimental condition: long photoperiod (16-hr light/8-hr dark at 18 c)"

ORIGIN

Query Match 36.9%; Score 483.2; DB 8; Length 895;
Best Local Similarity 99.4%; Pred. No. 1.3e-140;
Matches 485; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

813 GCATACCGGCAATTGATTGCTACAGCAAGCGGAGTTGCCCGAGCGCGCCAC 872
|||||
895 GCATACCGGCAATTGATTGCTACAGCAAGCGGAGTTGCCCGAGCGCGCCAC 836
|||||
873 CCCGTTATTAGATTGATCAAGACGCGTTCAGCCCATTCACCGCAAAAACAGCGTA 932
|||||
835 CCCGTTATTAGATTGATCAAGACGCGTTCAGCCCATTCACCGCAAAAACAGCGTA 776
|||||
933 TGGTGTGACATTACCACTTCAGTGTGTTTATCGCCGACACGATTAATCTGGCAAA 992
|||||
775 TGGTGTGACATTACCACTTCAGTGTGTTTATCGCCGACACGATTAATCTGGCAAA 716
|||||
993 TCTCGGCGGCGACCTGAGACCTGACGCTTCCCGGTTCAGCCGGAATAACCGCCGC 1052
|||||
715 TCTCGGCGGCGACCTGAGACCTGACGCTTCCCGGTTCAGCCGGAATAACCGCCGC 656
|||||
1053 AGGTGTGAACTGTGTGTTGAAAGCGTGGCGCTGAGCTAACGATTAACAGCCAGTGAATCA 1112
|||||
655 AGGTGTGAACTGTGTGTTGAAAGCGTGGCGCTGAGCTAACGATTAACAGCCAGTGAATCA 596
|||||
1113 GGTTCGCTGCTCTTCACAGCTTTACAGAGATGCGGTATTAACAGCCGCTGTCAATTA 1172
|||||
595 GGTTCGCTGCTCTTCACAGCTTTACAGAGATGCGGTATTAACAGCCGCTGTCAATTA 536
|||||
1173 TACGCGCGCGGAGAGGTGAACCTGACCTGGCAGGATGTGAAGAGCAAAATGCCGAGCG 1232

|||||
535 TACGCGCGCGGAGAGGTGAACCTGACAGATTTGAAGACGCAAAATGCCGAGCG 476
|||||
1233 CATTGTTGCTTGGCAGGTTTATACGCAATTCGTAATGAAGACGATACCGCGTGCAG 1292
|||||
475 CATTGTTGCTTGGCAGGTTTATACGCAATTCGTAATGAAGACGATACCGCGTGCAG 416
|||||

1293 TTGTGAT 1300
|||||
415 TTGTGAT 408

RESULT 7
CN754382/c 868 bp mRNA linear EST 19-MAY-2004
LOCUS ID0AA12DE01RM1 APMS Acyrthosiphon pisum cDNA clone ID0AA12DE01
DEFINITION 5', mRNA sequence.
ACCESSION CN754382
VERSION CN754382.1 GI:47519379
KEYWORDS EST.
SOURCE Acyrthosiphon pisum (pea aphid)
ORGANISM Acyrthosiphon pisum
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes; Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.

1 (bases 1 to 868)
Hunter, W., Martinez-Torres, D., Rabhe, Y., Sabater-Munoz, B., Stern, D., Tagu, D. and Wincker, P.
An expressed sequence tags database for the pea aphid *Acyrthosiphon pisum*

Unpublished (2004)
Contact: D. Tagu

JOURNAL
INRA Rennes
UMR BIO3P, BP 35327, F-35653 Le Rheu Cedex France
Tel: +33.2.23.48.51.65
Fax: +33.2.23.48.51.50
Risk of contamination by bacterial sequences from obligatory (Buchnera) or facultative endosymbionts. These sequences were obtained in the frame of the International Consortium of Aphid Genomics in collaboration with Genoscope

PCR Primers
FORWARD: CAGGAACGCTATGACC
Plate: 12

row: B column: 1.
Location/Qualifiers
1. 868
/organism="Acyrthosiphon pisum"
/mol_type="mRNA"
/cultivar="developmentstage"
/db_xref="taxon:7029"
/clone="ID0AA12DE01"
/tissue_type="whole insect"
/dev_stage="nymphs and adults (parthenogenetic females)"
/lab_host="XLI-Blue"
/clone_1lb="ApMS"
/note="Vector: pBS-SK minus, Site 1: EcoRI; Site 2: XhoI; Sample name: ID0AA1; Plant growth place: Department of Ecology & Evolutionary Biology, Princeton University; Soil conditions: Soil; Sowing date: 01/06/1999; Harvesting date: 01/06/1999; Stress date: no stress; Description: Aphids inoculated on one-week old *Vicia faba* under non-sterile conditions. All parthenogenetic stages and both winged and wingless adults were collected for library construction. ; experimental condition: long photoperiod (16-hr light/8-hr dark at 18 c)"

ORIGIN

Query Match 34.2%; Score 446.8; DB 8; Length 868;
Best Local Similarity 99.4%; Pred. No. 4.2e-123;
Matches 459; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

839 AACGACCCGAGAGTTCCCGCAGCCGCGCACCCGTTATTAGATTGATCAAGACAG 898
|||||
868 AACGACCCGAGAGTTCCCGCAGCCGCGCACCCGTTATTAGATTGATCAAGACAG 810

Qy 899 CGTTGACGCCCATCCACCGGAAAAAGAGGGGTATGTTGACATTACCACTTCAGAGC 958
| | | | |
Db 809 CTTTGACGCCCATCCATCCACCGGAAAAAGAGGGGTATGTTGACATTACCACTTCAGAGC 750
| | | | |
Qy 959 TGTATTATCGCCGAGCAGCATATCTATCTGGCAATCTCGCGCGCAGCTGAGCTCAACT 1018
| | | | |
Db 749 TGTATTATCGCCGAGCAGCATATCTATCTGGCAATCTCGCGCGCAGCTGAGCTCAACT 690
| | | | |
Qy 1019 GAGAGCTTCGCCGTGATAGCCGGATTAACAGCCGCCAGGTGTGAATGTGTGTTGAAGCT 1078
| | | | |
Db 689 GAGAGCTTCGCCGTGATAGCCGGATTAACAGCCGCCAGGTGTGAATGTGTGTTGAAGCT 630
| | | | |
Qy 1079 GAGCGTCGCTAAGCATTAACAGCCGAGTGAATTCAGTTTCGTGTCTTCCAGACTTAC 1138
| | | | |
Db 629 GAGCGTCGCTAAGCATTAACAGCCGAGTGAATTCAGTTTCGTGTCTTCCAGACTTAC 570
| | | | |
Qy 1139 AGCAGATCGTGAATTAACAGCCGCTGTCAATTAATACCCGCCGAGAGGTGAATCTGA 1198
| | | | |
Db 569 AGCAGATCGTGAATTAACAGCCGCTGTCAATTAATACCCGCCGAGAGGTGAATCTGA 510
| | | | |
Qy 1199 CCCTGGCAGATGTGAAGAGGAAATGCGCAGGGCATGTCTTCTGGCAGGTTTACCC 1258
| | | | |
Db 509 CCCTGGCAGATGTGAAGAGGAAATGCGCAGGGCATGTCTTCTGGCAGGTTTACCC 450
| | | | |
Qy 1259 AATATGTAATGAAGCAGCATACCGGCGTGCAGTTGAGAT 1300
| | | | |
Db 449 AATATGTAATGAAGCAGCATACCGGCGTGCAGTTGAT 408
| | | | |

RESULT 8
CN757565/c 864 bp mRNA linear EST 20-MAY-2004
LOCUS ID0AAA1DP09RM1 ApMS Acyrthosiphon pisum cDNA clone ID0AAA1DP09 5',
DEFINITION mRNA sequence.
ACCESSION CN757565
VERSION CN757565.1 GI:47531488
KEYWORDS EST.
SOURCE Acyrthosiphon pisum (pea aphid)
ORGANISM Acyrthosiphon pisum
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.
REFERENCE 1 (bases 1 to 864)
AUTHORS Hunter W., Martinez-Torres D., Rahbe Y., Sabater-Munoz B.,
Stern D., Tagu D. and Wincker P.
TITLE An expressed sequence tags database for the pea aphid Acyrthosiphon
pisum
JOURNAL Unpublished (2004)
COMMENT Contact: D. Tagu
INRA Rennes
UMR BIO3P, BP 35327, F-35653 Le Rheu Cedex France
Tel: +33.2.23.48.51.65
Fax: +33.2.23.48.51.50
Risk of contamination by bacterial sequences from obligatory
(Buchnera) or facultative endosymbionts. These sequences were
obtained in the frame of the International Consortium of Aphid
Genomics in collaboration with Genoscope
PCR Primers
FORWARD: CAGGAAACAGCTATGACC
Plate: 1 row: F column: 9.
Location/Qualifiers
1..864
/organism="Acyrthosiphon pisum"
/mol_type="mRNA"
/cultivar="developmentstage"
/db_xref="taxon:7029"
/clone="ID0AAA1DP09"
/tissue_type="whole insect"
/dev_stage="nymphs and adults (parthenogenetic females)"
/lab_host="XLI-Blue"
/clone_lib="ApMS"
/note="Vector: pBS-SK minus; Site_1: EcoRI; Site_2: XhoI;

Sample name: ID0AAA ; Plant growth place: Department of
Ecology & Evolutionary Biology, Princeton University ;
Soil conditions: Soil ; Sowing date: 01/06/1999 ;
Harvesting date: 01/06/1999 ; Stress date: no stress ;
Description: Aphids inoculated on one-week old *Vicia faba*
under non-sterile conditions. All parthenogenetic stages
and both winged and wingless adults were collected for
library construction. ; experimental condition: long
photoperiod (16-hr light/8-hr dark at 18 c°)

ORIGIN

Query Match 33.9% ; Score 442.8 ; DB 8 ; Length 864 ;
Best Local Similarity 99.3% ; Pred. No. 7.7e-128 ;
Matches 455 ; Conservative 0 ; Mismatches 2 ; Indels 1 ; Gaps 1 ;

Qy 843 CAGCCAGAGGTTCGCCGACCCGCCACCCCTTATTTAGATTATCAAGAGAGGTT 902
| | | | |
Db 864 CAGCCAGAGGTTCGCCGACCCGCCACCCCTTATTTAGATTATCAAGAGAGGTT 806
| | | | |
Qy 903 GAGGCCCATCAGCCGAAAAAGAGCGTATGTTGATGATTAACCACTTCAGTGTGT 962
| | | | |
Db 805 GAGGCCCATCAGCCGAAAAAGAGCGTATGTTGATGATTAACCACTTCAGTGTGT 746
| | | | |
Qy 963 TATCGCCGACACGATATCTATCTGGCAATCTCGCGCGCACTGAGACTCACTGAGC 1022
| | | | |
Db 745 TATCGCCGACACGATATCTATCTGGCAATCTCGCGCGCACTGAGACTCACTGAGC 686
| | | | |
Qy 1023 GCTTCCCGGTACGCCGATTAACAGCCGCGAGTGTGAATCTGTGTTGAACCTGGCG 1082
| | | | |
Db 685 GCTTCCCGGTACGCCGATTAACAGCCGCGAGTGTGAATCTGTGTTGAACCTGGCG 626
| | | | |
Qy 1083 TCGGCTAAGCATTAACAGCCAGTGAATTCAGTTTCGTCTTCAGACTTTACAGCA 1142
| | | | |
Db 625 TCGGCTAAGCATTAACAGCCAGTGAATTCAGTTTCGTCTTCAGACTTTACAGCA 566
| | | | |
Qy 1143 GATGCGTATTAACAGCCGCTGTCAATTAATACCCGCCGAGAGGTGAATCTGACCT 1202
| | | | |
Db 565 GATGCGTATTAACAGCCGCTGTCAATTAATACCCGCCGAGAGGTGAATCTGACCT 506
| | | | |
Qy 1203 GGCAGATGTGAAGAGGAAATGCGCAGGCGATGTTCTGTTGCAAGTTTACGAAT 1262
| | | | |
Db 505 GGCAGATGTGAAGAGGAAATGCGCAGGCGATGTTCTGTTGCAAGTTTACGAAT 446
| | | | |
Qy 1263 CGTGAATGAAGCAGCATACCGCGGTGCAAGTTGAGAT 1300
| | | | |
Db 445 CGTGAATGAAGCAGCATACCGCGGTGCAAGTTGAT 408
| | | | |

RESULT 9
BE520240/c 393 bp mRNA linear EST 19-MAR-2001
LOCUS M1B1S2TM Arabidopsis developing seed Arabidopsis thaliana cDNA
DEFINITION clone M1B12 5', mRNA sequence.
ACCESSION BE520240
VERSION BE520240.1 GI:9778242
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 393)
AUTHORS White J.A., Todd J., Newman T., Rocks N., Giralte T., Martinez de
Ibarra O., Javoritski J.G., Ohlrogge J. and Benning C.
TITLE A new set of Arabidopsis expressed sequence tags from developing
seeds. The metabolic pathway from carbohydrates to seed oil
JOURNAL Plant Physiol. 124 (4), 1582-1594 (2000)
PUBMED 1115876
COMMENT Contact: Benning, C
Dept. of Biochemistry & Molecular Biology
Michigan State University
224 Biochemistry, Michigan State University, East Lansing, MI
48824, USA

Tel: 517 355 1609
 Fax: 517 353 9334
 Email: benning@osu.edu
 Michigan State University DNA Sequencing Facility Arabidopsis
 Biological Resource Center, The Ohio State University, 309 Botany &
 Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210 USA, FAX:
 6142920603 TBI: 6142929371.

FEATURES

source
 1..393
 Location/Qualifiers
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /ecotype="Columbia"
 /db_xref="taxon:3702"
 /clone="M1812"
 /issue_type="seed"
 /dev_stage="5-13 days after flowering"
 /lab_host="E.coli"
 /clone_1lb="Arabidopsis developing seed"
 /note="Organ: Developing seed; Vector: pBluescript SK-;
 Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 28.5%; Score 373.4; DB 7; Length 393;
 Best Local Similarity 96.7%; Pred. No. 5.4e-106;
 Matches 380; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 319 ATTATTCGATGTGACGAGCGTACCCGTAAACAGCGGAAGCTTCGCCGCCGCGCTG 378
 Db 393 ATTATTCGATGTGACGAGCGTACCCGTAAACAGCGGAAGCTTCGCCGCCGCGCTG 334
 QY 379 GCACCTGACGTGACATACCGTACATCCGAGGAGATATCCAGTCCGATCCGCTTA 438
 Db 333 GCACCTGACGTGACATACCGTACATCCGAGGAGATATCCAGTCCGATCCGCTTA 274
 QY 439 TTTAATCCTCTTAAAACTGGCGTTTCCCACTGATACGCAACGTGACTGACGCGATC 498
 Db 273 TTTAATCCTCTTAAAACTGGCGTTTCCCACTGATACGCAACGTGACTGACGCGATC 214
 QY 499 CTCGAGAGGCGAGAGGCTCAATTGCTGACTTTACCGGCGATTTATCAACCGCGTTTCCG 558
 Db 213 CTCGAGAGGCGAGAGGCTCAATTGCTGACTTTACCGGCGATTTATCAACCGCGTTTCCG 154
 QY 559 GAACCTGGAACGGGCTGTAATTTCCGCAATCAAACTTGCGCTTAAACGTGAGAAACG 618
 Db 153 GAACCTGGAACGGGCTGTAATTTCCGCAATCAAACTTGCGCTTAAACGTGAGAAACG 94
 QY 619 GACGAAGCTGTCTTAAACGAGGATTCATCGGAATCAAGGTGAGCCCGCATCT 678
 Db 93 GACGAAGCTGTCTTAAACGAGGATTCATCGGAATCAAGGTGAGCCCGCATCT 34
 QY 679 GTCTCATTAACCGGTGCGGTAAAGCTCGCATCA 711
 Db 33 GTCTCATTAACCGGTGCGGTAAAGCTCGCATCA 1

RESULT 10

AM036132 383 bp mRNA linear EST 18-MAY-2001
 LOCUS EST274508 tomato seed, TMMU Lycopersicon esculentum cDNA clone
 DEFINITION CLEB1R23 similar to periplasmic phosphoanhydride phosphohydrolase
 precursor, putative, mRNA sequence.

ACCESSION

AM036132.1 GI:5894811

VERSION

AM036132.1 GI:5894811
 EST.
 Lycopersicon esculentum (Solanum lycopersicum)

KEYWORDS

Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE

Alcala, J., Vrebalov, J., White, R., Matern, A. L., Vision, T.,
 Holt, I. E., Liang, F., Upcon, J., Ronning, C. M., Craven, M. B.,
 Fujii, C. Y., Bowman, C. L., Nierman, W., Frazer, C. M., Venter, J. C.,

TITLE
 JOURNAL
 COMMENT
 Martin G. B., Tankley, S. D. and Giovannoni, J.
 Generation of ESTs from tomato seed tissue
 Unpublished (1999)
 Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 3 prime sequence.

FEATURES

source
 1..383
 Location/Qualifiers
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /culivar="TA496"
 /db_xref="taxon:4081"
 /clone="CLEB1R23"
 /issue_type="seeds"
 /dev_stage="quiescent seed"
 /lab_host="XLI-Blue WRF"
 /clone_1lb="tomato seed, TMMU"
 /note="Vector: pBluescript SK(-); Site_1: EcoRI, Site_2:
 XhoI; cDNA - Tomato Seed EST Library. Directionally cloned
 cDNAs inserted into pBluescript SK(-) at 5' end with
 EcoRI and 3' end with XhoI site."

ORIGIN

Query Match 28.3%; Score 370.2; DB 7; Length 383;
 Best Local Similarity 97.9%; Pred. No. 5.5e-105;
 Matches 375; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 348 TAAACAGGCGAAGCTTCGCCGCCGCGGCTGCGACCTGCACTGCAATTAACGTCATAC 407
 Db 1 TAAACAGGCGAAGCTTCGCCGCCGCGGCTGCGACCTGCACTGCAATTAACGTCATAC 60
 QY 408 CCAGGCAAGTACGTCCAGTCCGATCCGTTATTAATCCTCTAATAAAGTGGGTTGCCA 467
 Db 61 CCAGGCAAGTACGTCCAGTCCGATCCGTTATTAATCCTCTAATAAAGTGGGTTGCCA 120
 QY 468 ACTGATTAACGGAAAGTGAAGTACGATCTCTGAGAGGCGAGGAGGTCATTTGCTGA 527
 Db 121 ACTGATTAACGGAAAGTGAAGTACGATCTCTGAGAGGCGAGGAGGTCATTTGCTGA 180
 QY 528 CTTTACCGGCGCATTAATCAACGGCGCTTCGGAATCGGAACGGGCTTAATTTCCGCA 587
 Db 191 CTTTACCGGCGCATTAATCAACGGCGCTTCGGAATCGGAACGGGCTTAATTTCCGCA 240
 QY 588 ATCAACTTGTGCTTAAACGTGAGAAACAGGACGAAGCTGTCTTAAACGAGGACTT 647
 Db 241 ATCAACTTGTGCTTAAACGTGAGAAACAGGACGAAGCTGTCTTAAACGAGGACTT 300
 QY 648 ACCATGGAATCAAGGTGAGGCGGACTGTGTCTTAAACGGGTGCGTAAGCTTCG 707
 Db 301 ACCATGGAATCAAGGTGAGGCGGACTGTGTCTTAAACGGGTGCGTAAGCTTCG 360
 QY 708 ATCAATGCTGACGGAATATTTTC 730
 Db 361 ATCAATGCTGACGGAATATTTTC 383

RESULT 11

AM036134 383 bp mRNA linear EST 18-MAY-2001
 LOCUS EST274510 tomato seed, TMMU Lycopersicon esculentum cDNA clone
 DEFINITION CLEB1R23 similar to periplasmic phosphoanhydride phosphohydrolase
 precursor, putative, mRNA sequence.

ACCESSION

AM036134.1 GI:5894813

VERSION

AM036134.1 GI:5894813
 EST.
 Lycopersicon esculentum (Solanum lycopersicum)

KEYWORDS

Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 383)
AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.L., Viston,T.,
Holt,I.B., Liang,P., Upton,J., Konning,C.M., Craven,M.B.,
Fujii,C.Y., Bowman,C.L., Nierman,W., Frieser,C.M., Venter,J.C.,
Martin,G.B., Tanksley,S.D. and Giovannoni,J.
TITLE Generation of ESTs from tomato seed tissue
JOURNAL Unpublished (1999)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.

FEATURES
source
location/Qualifiers
1..383
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultiivar="TA96"
/db_xref="taxon:4081"
/clone="CLEB1E23"
/tissue_type="seeds"
/dev_stage="quiescent seed"
/lab_host="XLI-Blue MRP"
/clone_lib="tomato seed, TAMU"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; CLER - Tomato Seed EST library. Directionally cloned
cDNAs inserted into pBluescript SK(-) at 5' end with
EcoRI and 3' end with XhoI site."

ORIGIN
Query Match 28.3%; Score 370.2; DB 7; Length 383;
Best Local Similarity 97.9%; Pred. No. 5.5e-105;
Matches 375; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 348 TAAACAGCGCAAGCCTTCGCCGCCGCTGCACTGTGCAATACCGTACATAC 407
Db 383 TAAACAGCGCAAGCCTTCGCCGCCGCTGCACTGTGCAATACCGTACATAC 324
Qy 408 CCAGGCAATACGTCACATCCGATCCGTTATTATCTCTAATAACCTGCGCTTGGCA 467
Db 323 CCAGGCAATACGTCACATCCGATCCGTTATTATCTCTAATAACCTGCGCTTGGCA 264
Qy 468 ACTGGATACCGGAACGTCACGCGATCCTCGAGAGGGGCAAGGGGCAATTGCGCA 527
Db 263 ACTGGATACCGGAACGTCACGCGATCCTCGAGAGGGGCAAGGGGCAATTGCGCA 204
Qy 528 CTTTACCGGGCATTTATCAACCGCGCTTTCGGAACCTGGAACGGGTCTTAATTTTCCGCA 587
Db 203 CTTTACCGGGCATTCGGAACCGCGCTTTCGGAACCTGGAACGGGTCTTAATTTTCCGCA 144
Qy 588 ATCAACTTGTGCTTAAACGTGAGAAACAGAGCAAAAGCTGTTATTACGACGCAATT 647
Db 143 ATCAACTTGTGCTTAAACGTGAGAAACAGAGCAAAAGCTGTTATTACGACGCAATT 84
Qy 648 ACCATCGGAACTCAGAGTGAGCGCGCACTGTCTCATTAACCGGTGCGGTAAAGCTGCG 707
Db 83 ACCATCGGAACTCAGAGTGAGCGCGCAATGTCTCATTAACCGGTGCGGTAAAGCTGCG 24
Qy 708 ATCAATGCTGACGAGATATTTC 730
Db 23 ATCAATGCTGACGAGATATTTC 1

RESULT 12
LOCUS BU713769 531 bp mRNA linear EST 23-OCT-2003
DEFINITION SUSAHBG01 Adult SJC 7/94 Schistosoma japonicum cDNA similar to
sp|P07102|PPA ECOLI PERIPLASMIC APPA PROTEIN PRECURSOR [INCLUDS:
PHOSPHOANTHRIDIDE PHOSPHOHYDROLASE (PH 2.5 ACID PHOSPHATASE) (AP);
6-PHYTASE], mRNA sequence.
ACCESSION BU713769
KEYWORDS EST.

SOURCE
ORGANISM Schistosoma japonicum
REFERENCE Schistosoma japonicum
AUTHORS Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
Strigoida; Schistosomatoidae; Schistosomatidae; Schistosoma.
1 (bases 1 to 531)
Hu,W., Yan,Q., Shen,D.K., Liu,F., Zhu,Z.D., Song,H.D., Xu,X.R.,
Wang,Z.J., Rong,Y.P., Zeng,L.C., Wu,J., Zhang,X., Wang,J.J.,
Xu,X.N., Wang,S.Y., Fu,G., Zhang,X.L., Wang,Z.Q., Brindley,P.J.,
McManus,B.P., Xue,C.L., Feng,Z., Chen,Z., and Han,Z.G.
TITLE Evolutionary and biomedical implications of a Schistosoma japonicum
JOURNAL complementary DNA resource
PUBMED Nat. Genet. 35 (2), 139-147 (2003)
12973349
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn.

FEATURES
source
location/Qualifiers
1..531
/organism="Schistosoma japonicum"
/mol_type="mRNA"
/strain="Chinese (Anhui) strain"
/db_xref="taxon:6182"
/sex="Male and female"
/tissue_type="Whole body"
/dev_stage="Adult worms"
/lab_host="Mouse and rabbit"
/clone_lib="Adult SJC 7/94"
/note="Vector: Lambda ZAP-II XR; Site_1: EcoR I; Site_2:
XhoI; Several hundred adult Schistosoma japonicum
(Anhui, P.R. China, strain), of mixed sex, were perfused
from the mesenteries of experimentally infected mice and
rabbits at the Queensland Institute of Medical Research,
Brisbane, Australia (QIMR), and stored for several months
in liquid nitrogen. Subsequently, mRNA was isolated at the
QIMR from lysates of these worms by oligo dT
chromatography, using a kit from Pharmacia. The mRNA was
then shipped to Clontech, Palo Alto, CA, USA, who
constructed a cDNA library. First strand synthesis was
primed with an oligo-dT-XhoI-primer and synthesized using
M-MuLV reverse transcriptase. Second strand synthesis was
accomplished with RNase H and T4 DNA polymerase. The
double stranded cDNA was ligated to EcoRI linkers,
digested with EcoRI and XhoI, and ligated into the
phagemid vector lambda ZAP II XR. After construction of
this directional library by Clontech, it was returned to
the QIMR. During analysis of the library at the QIMR, we
have found that a small percentage, 2% to 3%, of the
clones contain inserts that appear to be highly homologous
to sequences from salmonoid fishes, as determined by
homology comparisons using BLAST and by Southern
hybridization analysis to genomic DNA from salmon (Sigma
Chemical Co., St. Louis, MO) under stringent washing
conditions. The remainder of the clones appear to contain
S. japonicum sequences."

ORIGIN
Query Match 26.7%; Score 348.6; DB 3; Length 531;
Best Local Similarity 88.0%; Pred. No. 4.1e-99;
Matches 425; Conservative 0; Mismatches 54; Indels 4; Gaps 4;

Qy 670 GCCGAGTGTCTCATTAACCGGTGCGGTAAAGCTCCGATCAATGTCAGAGATATTT 729
Db 1 GCGACATGTCTATTAAACCGGTGCGGTAAAGCTCCGATCAATGTCAGAGATGGG 60
Qy 730 CTCTCGCAACAGCAGGGAATCCGAGCCGGGGTGGGGAAGATCCAGATTCAAC 789
Db 61 CTCTCGCAACAGCAGGGAATCCGAGCCGGGGTGGGGAAGATCCAGATTCAAC 120
Qy 790 CAGTGAACACCTTGCTAAGTTTGCAATACGGCGCAATTGATTGCTACAAACGACCCCA 849

ORIGIN /clone_1lib="Human Bone Marrow Stromal Fibroblast"
/note="Vector: pBluescript; Site_1: EcoRI; Site_2: XhoI"

Query Match 22.6%; Score 295.4; DB 1; Length 354;
Best Local Similarity 91.7%; Pred. No. 2.4e-81;
Matches 322; Conservative 0; Mismatches 28; Indels 1; Gaps 1;

QY 499 CTCGAGAGGGGAGAGGGTCAATTGCTTACCTTACCCGGGCATTTATCAAGCCGTTTCG 558
DB 350 CTTGACGAGGGGACAGAGGTCAATTGCTTATTTACCGGGCATCGCAACGGGCGTTTCG 291
QY 559 GAATCGAAGCGGCTTAAATTTTCCGCATCAACTTGTGCTTAAAGTGAAGAACAG 618
DB 290 GAATCGAAGCGGCTTAAATTTTCCGCATCAACTTGTGCTTAAAGTGAAGAACAG 232
QY 619 GACGAAAGCTGTTCATTAAACGAGGCAATTACCATCGAACTCAAGGTGAGCGCGGACTGT 678
DB 231 GACGAAAGCTGTTCATTAAACGAGGCAATTACCATCGAACTCAAGGTGAGCGCGGACTGT 172
QY 679 GTCTCATTAACCGGCGGCTTAAAGCTGACATCAATGCGAGGAGATATTTCTCCTGCA 738
DB 171 GTCTCATTAACCGGCGGCTTAAAGCTGACATCAATGCGAGGAGATATTTCTCCTGCA 112
QY 739 CAAGCACAGGAGATCCGAGCCGGGCTGGGGAAGATCAACGATTCACACAGTGAAC 798
DB 111 CAAGCACAGGAGATCCGAGCCGGGCTGGGGAAGATCAACGATTCACACAGTGAAC 52
QY 799 ACCTTGCTAAGTTTGCATTAACGCGCAATTGATTGCTTACACGACGCCA 849
DB 51 ACCTTGCTAAGTTTGCATTAACGCGCAATTGATTGCTTACACGACGCCA 1

RESULT 15

BU619443 707 bp mRNA linear EST 01-OCT-2003
LOCUS BU619443 NIBB Mochii normalized Xenopus early gastrula library
DEFINITION Xenopus laevis cDNA clone XLA189d24 5', mRNA sequence.

ACCESSION BU619443
VERSION BU619443.1 GI:37258203

KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis

REFERENCE 1 (bases 1 to 707)
AUTHORS Kohayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-i, T. and Kohara, Y.

TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-i

Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855

Email: tshini@gens.nig.ac.jp
The information of this clone is available through the following URL.

http://xenopus.nibb.ac.jp.
Location/Qualifiers

FEATURES
SOURCE 1. 707

/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="XLA189d24"
/tissue_type="whole embryo"
/dev_stage="stage 10.5"
/clone_lib="NIBB Mochii normalized Xenopus early gastrula library"

ORIGIN

Query Match 19.4%; Score 254.4; DB 2; Length 707;
Best Local Similarity 99.6%; Pred. No. 2.6e-68;
Matches 255; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 750 AATGCCGAGCCGGGGGTGGGAGAGATCACCGATTTCACACAGTGAACACCTTGCTTAAG 809
DB 1 AATGCCGAGCCGGGGGTGGGAGAGATCACCGATTTCACACAGTGAACACCTTGCTTAAG 60
QY 810 TTTCGATTAAGCGCAATTTGATTGCTTACAAAGCAGCAGAGGTTCCCGCAGCCGCGC 869
DB 61 TTTCGATTAAGCGCAATTTGATTGCTTACAAAGCAGCAGAGGTTCCCGCAGCCGCGC 120
QY 870 CACCCGTTATTTAGATTGATCAAGACAGGCTTGACGCCCCATCCACGCAAAAACAGGC 929
DB 121 CACCCGTTATTTAGATTGATCAAGACAGGCTTGACGCCCCATCCACGCAAAAACAGGC 180
QY 930 GTATGTTGTGACATTACCCACTTCAAGTCTGTTTATGCGCGGACACGATTAATCTGGC 989
DB 181 GTATGTTGTGACATTACCCACTTCAAGTCTGTTTATGCGCGGACACGATTAATCTGGC 240
QY 990 AATCTCGGCGGCGCA 1005
DB 241 AATCTCGGCGGCGCA 256

Search completed: June 14, 2006, 14:21:52
Job time : 6872.62 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocoelation Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2006, 10:29:05 ; Search time 1596.84 Seconds
(without alignments)
10064.996 Million cell updates/sec

Title: US-10-601-319-9

Perfect score: 1308

Sequence: 1 atgaagagcactatcc.....gcagttgagatcctatcta 1308

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main:

1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq.*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09_PUBCOMB.seq.*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq.*
15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq.*
16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1308	100.0	1308	US-10-156-660-1	Sequence 1, Appl1
2	1308	100.0	1308	US-10-601-319-9	Sequence 9, Appl1
3	1308	100.0	1308	US-10-933-115-9	Sequence 9, Appl1
4	1296	99.1	1296	US-11-056-354-1	Sequence 1, Appl1
5	1268	97.0	1901	US-10-601-319-6	Sequence 6, Appl1
6	1267	96.9	1901	US-10-601-319-5	Sequence 5, Appl1
7	1267	96.9	1901	US-10-933-115-5	Sequence 5, Appl1
8	1265	96.7	1323	US-09-777-556A-1	Sequence 1, Appl1
9	1265	96.7	1323	US-09-866-379-1	Sequence 1, Appl1
10	1265	96.7	1323	US-10-034-985-1	Sequence 1, Appl1
11	1265	96.7	1323	US-10-430-355-1	Sequence 1, Appl1
12	1265	96.7	1323	US-10-601-319-1	Sequence 1, Appl1
13	1265	96.7	1323	US-10-933-115-1	Sequence 1, Appl1
14	1264	96.6	1299	US-10-282-122A-7167	Sequence 7167, Ap
15	1263.8	96.6	1901	US-09-866-379-6	Sequence 6, Appl1
16	1263.8	96.6	1901	US-09-866-379-7	Sequence 7, Appl1
17	1263.8	96.6	1901	US-10-156-660-3	Sequence 3, Appl1

18	1263.8	96.6	1901	8	US-10-601-319-7	Sequence 7, Appl1
19	1263.8	96.6	1901	10	US-10-933-115-6	Sequence 6, Appl1
20	1263.8	96.6	1901	10	US-10-933-115-7	Sequence 7, Appl1
21	1263.8	96.6	1901	14	US-11-056-354-3	Sequence 3, Appl1
22	1262.2	96.5	1901	3	US-09-866-379-5	Sequence 5, Appl1
23	1262.2	96.5	1901	3	US-09-866-379-9	Sequence 9, Appl1
24	1233.6	94.3	1486	13	US-11-018-709-2	Sequence 2, Appl1
25	1228	93.9	1489	6	US-10-266-041-9	Sequence 9, Appl1
26	1228	93.9	1489	7	US-10-284-962-1	Sequence 1, Appl1
27	1222.4	93.5	1486	7	US-10-284-962-4	Sequence 4, Appl1
28	1222.4	93.5	1486	13	US-11-018-709-4	Sequence 4, Appl1
29	758.2	58.0	11357	13	US-11-074-522-14	Sequence 14, Appl1
30	691.4	52.9	1281	7	US-10-334-672-4	Sequence 4, Appl1
31	691.4	52.9	1281	7	US-10-334-671-4	Sequence 4, Appl1
32	691	52.8	1737	10	US-10-450-763-25689	Sequence 25689, A
33	404.8	30.9	466	8	US-10-282-122A-3181	Sequence 3181, Ap
34	377.2	28.8	1264	10	US-10-450-763-24247	Sequence 24247, A
35	377.2	28.8	1264	10	US-10-450-763-30120	Sequence 30120, A
36	377	28.8	4050	10	US-10-450-763-386	Sequence 386, App
37	230.2	17.6	1326	8	US-10-282-122A-41608	Sequence 41608, A
38	226.2	17.3	1325	6	US-10-021-723A-3	Sequence 3, Appl1
39	223	17.0	1325	6	US-10-021-723A-11	Sequence 11, Appl1
40	212.4	16.2	1326	6	US-10-021-723A-11	Sequence 11, Appl1
41	191	14.6	356	10	US-10-450-763-25687	Sequence 25687, A
42	164.2	12.6	254	7	US-10-317-444-301	Sequence 301, App
43	164.2	12.6	254	7	US-10-317-444-302	Sequence 302, App
44	164.2	12.6	254	7	US-10-317-444-303	Sequence 303, App
45	164.2	12.6	254	7	US-10-317-444-304	Sequence 304, App

ALIGNMENTS

RESULT 1
US-10-156-660-1
Sequence 1, Application US/10156660
Publication No. US20030103958A1
GENERAL INFORMATION:
APPLICANT: Short, Jay M.
APPLICANT: Kretz, Keith
APPLICANT: Gray, Kevin A.
APPLICANT: Barton, Nelson R.
APPLICANT: Garrett, James B.
APPLICANT: O'Donoghue, Eileen
TITLE OF INVENTION: PHYLASES, NUCLEIC ACIDS ENCODING THEM
FILE REFERENCE: 09010-029007
CURRENT APPLICATION NUMBER: US/10/156,660
CURRENT FILING DATE: 2002-10-01
PRIOR APPLICATION NUMBER: US 09/866,379
PRIOR FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: US 09/580,515
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 09/318,528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: US 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 09/259,214
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: US 08/910,798
PRIOR FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1308
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: modified phytase enzyme
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1308)

US-10-156-660-1

Query Match 100.0%; Score 1308; DB 6; Length 1308;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGGAGATCTTAATCCATTTTATCTCTTCGTATCCGTTAAACCCGCAATCTGCA 60
DB 1 ATGAAGGAGATCTTAATCCATTTTATCTCTTCGTATCCGTTAAACCCGCAATCTGCA 60
QY 61 TTGCGTCAGAGTGAAGCCGGAAGCTGAAGTGGAAAGTGTGATTTGTCAATGTCATGCT 120
DB 61 TTGCGTCAGAGTGAAGCCGGAAGCTGAAGTGGAAAGTGTGATTTGTCAATGTCATGCT 120
QY 121 GTGCGTGTCCAAACCAAGGCCACGCAACTGATGCAAGATGTCACCCGCAAGCATGAGCA 180
DB 121 GTGCGTGTCCAAACCAAGGCCACGCAACTGATGCAAGATGTCACCCGCAAGCATGAGCA 180
QY 181 ACCTGGCCGGTAAACTGGGTGAGCTGACACCGCGCGGTGGTGAAGCTAATCGCTATCTC 240
DB 181 ACCTGGCCGGTAAACTGGGTGAGCTGACACCGCGCGGTGGTGAAGCTAATCGCTATCTC 240
QY 241 GGAATTAATCTGGCGCTCAGCGCTCTGATGAGCGGACGAGATTGCTCTTAAATGTCCTCCG 300
DB 241 GGAATTAATCTGGCGCTCAGCGCTCTGATGAGCGGACGAGATTGCTCTTAAATGTCCTCCG 300
QY 301 CAGTCTGGTCAAGTCCGGAATTAATGCTGATGTCAGACGAGCGTACCCGTTAAACAGGGGAA 360
DB 301 CAGTCTGGTCAAGTCCGGAATTAATGCTGATGTCAGACGAGCGTACCCGTTAAACAGGGGAA 360
QY 361 GCCTTCGCGCGCGGCTGGACCTGACTGTGCAATTAACCGTACATACCAGGAGATACG 420
DB 361 GCCTTCGCGCGCGGCTGGACCTGACTGTGCAATTAACCGTACATACCAGGAGATACG 420
QY 421 TCCGATCCCGATCCGTTATTTAACTCTTAAACCTGCGCTTTGCAATGCAATGACGCG 480
DB 421 TCCGATCCCGATCCGTTATTTAACTCTTAAACCTGCGCTTTGCAATGCAATGACGCG 480
QY 481 AAGGTGAGTGAAGGATCTCTGAGAGGGCAGAGGGTCAATGTCGACTTTAACGGGGAT 540
DB 481 AAGGTGAGTGAAGGATCTCTGAGAGGGCAGAGGGTCAATGTCGACTTTAACGGGGAT 540
QY 541 TATCAAAACGGGCTTTCGGAACCTGGAACGGGTGCTTAATTTCCGCAATCAAACTTGTGC 600
DB 541 TATCAAAACGGGCTTTCGGAACCTGGAACGGGTGCTTAATTTCCGCAATCAAACTTGTGC 600
QY 601 CTTTAAAGTGAAGAACAGGACGAAGCTGTTCAATTAACGAGCATTAACATCGGAACTC 660
DB 601 CTTTAAAGTGAAGAACAGGACGAAGCTGTTCAATTAACGAGCATTAACATCGGAACTC 660
QY 661 AAGGTGAGCGCGGACTGTGCTCATTTAACCGGTGCGGAAGCTCGACTCAATGCTGACG 720
DB 661 AAGGTGAGCGCGGACTGTGCTCATTTAACCGGTGCGGAAGCTCGACTCAATGCTGACG 720
QY 721 GAGATATTTCTCTGCAACAAGCAAGGGAATGCGAGCGCGGGGTGGGAAGATCAC 780
DB 721 GAGATATTTCTCTGCAACAAGCAAGGGAATGCGAGCGCGGGGTGGGAAGATCAC 780
QY 781 GATTCAACACCACTGGAACACTTGTCTAAGTTTGATTAACGGCAATTTGATTTCTTCAA 840
DB 781 GATTCAACACCACTGGAACACTTGTCTAAGTTTGATTAACGGCAATTTGATTTCTTCAA 840
QY 841 CGCAGCGCAAGGTTGCGCGGACGCGGCAACCGGTTATGATTTGATTAACAAGCGG 900
DB 841 CGCAGCGCAAGGTTGCGCGGACGCGGCAACCGGTTATGATTTGATTAACAAGCGG 900
QY 901 TTGAAGCCCATTCACCGCAAAAAACAGCGGTATGATGATTAACCATCTTCAGTGTG 960
DB 901 TTGAAGCCCATTCACCGCAAAAAACAGCGGTATGATGATTAACCATCTTCAGTGTG 960
QY 961 TTTATCGCGGACACGATTAATCTGCAAAATCTCGCGCGGCACTGAGACTCAACTGG 1020
DB 961 TTTATCGCGGACACGATTAATCTGCAAAATCTCGCGCGGCACTGAGACTCAACTGG 1020

QY 1021 ACCTTCCGGGTACAGCCGGAATTAACAGCCGCGCAGGTGTGAATGATGTTGAACGCTGG 1080
DB 1021 ACCTTCCGGGTACAGCCGGAATTAACAGCCGCGCAGGTGTGAATGATGTTGAACGCTGG 1080
QY 1081 CGTCGCTAAGCAGTAACAGCCAGTGAATTCAGGTTTGGTGTCTTCCAGACTTTACAG 1140
DB 1081 CGTCGCTAAGCAGTAACAGCCAGTGAATTCAGGTTTGGTGTCTTCCAGACTTTACAG 1140
QY 1141 CAGATGCTGATTAACAGCCGCTGTCTTAATATCGCCGCGGAGAGGTGAACCTGACC 1200
DB 1141 CAGATGCTGATTAACAGCCGCTGTCTTAATATCGCCGCGGAGAGGTGAACCTGACC 1200
QY 1201 CTGGCAGATGTGAAGCCGAAATGCGCAGGCAATGCTGTGCTGTCAGAGTTTACGCA 1260
DB 1201 CTGGCAGATGTGAAGCCGAAATGCGCAGGCAATGCTGTGCTGTCAGAGTTTACGCA 1260
QY 1261 ATCGTAATGAAGCAGCATACCGGCGTGCAGTTTGAATCTCATCTA 1308
DB 1261 ATCGTAATGAAGCAGCATACCGGCGTGCAGTTTGAATCTCATCTA 1308

RESULT 2

US-10-601-319-9
; Sequence 9, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
; TITLE OR INVENTION: AND USING THEM
; FILE REFERENCE: 09010-029011
; CURRENT APPLICATION NUMBER: US/10/601,319
; PRIOR FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1308
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: modified phytase enzyme
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1296)
US-10-601-319-9

Query Match 100.0%; Score 1308; DB 8; Length 1308;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGGAGATCTTAATCCATTTTATCTCTTCGTATCCGTTAAACCCGCAATCTGCA 60
DB 1 ATGAAGGAGATCTTAATCCATTTTATCTCTTCGTATCCGTTAAACCCGCAATCTGCA 60
QY 61 TTGCGTCAAGTGAAGCCGGAAGCTGAAGTGGAAAGTGTGATTTGTCAATGTCATGCT 120

```
Db      61 TTGCGCTCAGAGTGAAGCCGAGCTGAAGTGAAGTGTGTGATGTCAGTGTCAATGCT 120
Qy      121 GTGCGGTCTCCAAACCAAGCCCAACGCAATGATGCAAGATGTCAACCCAGAGCGATGCGCA 180
Db      121 GTGCGGTCTCCAAACCAAGCCCAACGCAATGATGCAAGATGTCAACCCAGAGCGATGCGCA 180
Qy      181 ACCGCGCGGTAAACCTGGGTGAGTGAACACCGCGCGGTGAGTGAAGTATGCTATCTC 240
Db      181 ACCGCGCGGTAAACCTGGGTGAGTGAACACCGCGCGGTGAGTGAAGTATGCTATCTC 240
Qy      241 GGACATTAATGCGCTCAGCGCTGTGTGATGCGGACGCAATGCTGCTAAATGTGCTGCGCG 300
Db      241 GGACATTAATGCGCTCAGCGCTGTGTGATGCGGACGCAATGCTGCTAAATGTGCTGCGCG 300
Qy      301 CAGTCTGTGTCAGGTGCGGATTTGTGTGATGTCGACGAGCGTAACCCGTAAACAGCGCA 360
Db      301 CAGTCTGTGTCAGGTGCGGATTTGTGTGATGTCGACGAGCGTAACCCGTAAACAGCGCA 360
Qy      361 GCCTTCGCGCGGCGTGGCACTGTGATTAACCTGATTAACCTGATTAACCTGATTAACCT 420
Db      361 GCCTTCGCGCGGCGTGGCACTGTGATTAACCTGATTAACCTGATTAACCTGATTAACCT 420
Qy      421 TCCAGTCCCGATCCGTTATTTATCTCTTAAACCTGCGTGGCACTGTGATTAACCGG 480
Db      421 TCCAGTCCCGATCCGTTATTTATCTCTTAAACCTGCGTGGCACTGTGATTAACCGG 480
Qy      481 AAGCGTACTGACGCGATCTCTGAGAGGCGAGAGGCTCAATGCTGATTTAACCGGCGAT 540
Db      481 AAGCGTACTGACGCGATCTCTGAGAGGCGAGAGGCTCAATGCTGATTTAACCGGCGAT 540
Qy      541 TATCAAGCGCGCTTTCGCGAATGGAACGGGTGCTTAATTTCCGCAATCAACTTGTGCG 600
Db      541 TATCAAGCGCGCTTTCGCGAATGGAACGGGTGCTTAATTTCCGCAATCAACTTGTGCG 600
Qy      601 CTTAAACGTGAGAAACAGGACGAAAGCTGTCATTAACGACGACATTAACATCGGAATC 660
Db      601 CTTAAACGTGAGAAACAGGACGAAAGCTGTCATTAACGACGACATTAACATCGGAATC 660
Qy      661 AAGGTAGCGCGGACGTGTCTCTATTAACCGGTGCGGTAAAGCTGCGATCAATGCTGAG 720
Db      661 AAGGTAGCGCGGACGTGTCTCTATTAACCGGTGCGGTAAAGCTGCGATCAATGCTGAG 720
Qy      721 GAGATATTTCTCTCTGCAACAAACAGGAAATGCGGAGCGCGGTGGGGAAGATCAAC 780
Db      721 GAGATATTTCTCTCTGCAACAAACAGGAAATGCGGAGCGCGGTGGGGAAGATCAAC 780
Qy      781 GATTCACACCAAGTGAACACTTGTGAATGTCATTAACGCGCAATTTGATTTGCTACAA 840
Db      781 GATTCACACCAAGTGAACACTTGTGAATGTCATTAACGCGCAATTTGATTTGCTACAA 840
Qy      841 CGCAGCGCAAGAGTGTGCGCGAGCGCGCACCCCGTTATTAAGATTGATCAAGACAGG 900
Db      841 CGCAGCGCAAGAGTGTGCGCGAGCGCGCACCCCGTTATTAAGATTGATCAAGACAGG 900
Qy      901 TTGAAGCGCCCATTCACCGGAAAAACAGGCGTATGCTGATTAACCACTTCACTGCTG 960
Db      901 TTGAAGCGCCCATTCACCGGAAAAACAGGCGTATGCTGATTAACCACTTCACTGCTG 960
Qy      961 TTTATCGCGGACAGATTAATCTGGAATCTCGGCGGCGCATGGAAGCTCAACTGCG 1020
Db      961 TTTATCGCGGACAGATTAATCTGGAATCTCGGCGGCGCATGGAAGCTCAACTGCG 1020
Qy      1021 ACGCTTCCCGGTACCGCGATTAACGCGCGCGAGGTGTGAATCTGTGTTGAACGCTGCG 1080
Db      1021 ACGCTTCCCGGTACCGCGATTAACGCGCGCGAGGTGTGAATCTGTGTTGAACGCTGCG 1080
Qy      1081 CGTGGCTTAAGGATTAACAGCGATGATTTGCGTGTCTTCCAGCTTTACG 1140
Db      1081 CGTGGCTTAAGGATTAACAGCGATGATTTGCGTGTCTTCCAGCTTTACG 1140
Qy      1141 CAGATGCGGATTAACCGCGCTGTCAATTAATACGCGCGCGGAGAGGTGAACCTGACC 1200
Db      1141 CAGATGCGGATTAACCGCGCTGTCAATTAATACGCGCGCGGAGAGGTGAACCTGACC 1200
```

```
Db      1141 CAGATGCGGATTAACCGCGCTGTCAATTAATACGCGCGCGGAGAGGTGAACCTGACC 1200
Qy      1201 CTGGCAGGATGTGAAGACCGAAATGCGGACATGTGTGTTGGCAGGTTTACGCA 1260
Db      1201 CTGGCAGGATGTGAAGACCGAAATGCGGACATGTGTGTTGGCAGGTTTACGCA 1260
Qy      1261 ATCGTAATGAAGACGAGATACCGGCGTGCAGTTTGAATCTCATCTA 1308
Db      1261 ATCGTAATGAAGACGAGATACCGGCGTGCAGTTTGAATCTCATCTA 1308

RESULT 3
US-10-933-115-9
; Sequence 9, Application US/10933115
; Publication No. US20050281792A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES
; FILE REFERENCE: 564462001822
; CURRENT APPLICATION NUMBER: US/10/933,115
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1308
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: modified phytase enzyme
; NAME/KEY: CDS
; LOCATION: (1)...(1296)
US-10-933-115-9

Query Match      100.0%; Score 1308; DB 10; Length 1308;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 241 GGAATTACTGCGCTGAGCGCTGCTGAGTACCGAGGATTGCTGCTAAATGTGCTGCCG 300
Db 241 GGAATTACTGCGCTGAGCGCTGCTGAGTACCGAGGATTGCTGCTAAATGTGCTGCCG 300
Qy 301 CAGTCTGCTGAGTGTGCGATTAATGCTGATGTGCAAGAGCGTACCCGTAACAGCGGAA 360
Db 301 CAGTCTGCTGAGTGTGCGATTAATGCTGATGTGCAAGAGCGTACCCGTAACAGCGGAA 360
Qy 361 GCTTTCGCGCGCGGCTGCGACCTGACCTGCAATTAACCGTAACATCCAGGAGATTCG 420
Db 361 GCTTTCGCGCGCGGCTGCGACCTGACCTGCAATTAACCGTAACATCCAGGAGATTCG 420
Qy 421 TCCAGTCCCGATCCGTTATTAATCTCTAAACCTGCGCTTGGCACTGATTAACGCG 480
Db 421 TCCAGTCCCGATCCGTTATTAATCTCTAAACCTGCGCTTGGCACTGATTAACGCG 480
Qy 481 AACGTGACGCGGATCTTGAAGAGGCGAGAGGCTCAATTCGATTTACCGGCGAT 540
Db 481 AACGTGACGCGGATCTTGAAGAGGCGAGAGGCTCAATTCGATTTACCGGCGAT 540
Qy 541 TATCAACGCGGCTTTCGGAAGTGAACGCGGCTTAATTTCCGCAATCAACTGTGC 600
Db 541 TATCAACGCGGCTTTCGGAAGTGAACGCGGCTTAATTTCCGCAATCAACTGTGC 600
Qy 601 CTTAAACGTGAAGAACAGAGCAAAAGCTGTTCAATTAACGAGGATTAACATCCGAACTC 660
Db 601 CTTAAACGTGAAGAACAGAGCAAAAGCTGTTCAATTAACGAGGATTAACATCCGAACTC 660
Qy 661 AAGGTAGCGCGGATGTGTCTCATTAACGCGTGGGTGAAGCTTCGCAATGCTGACG 720
Db 661 AAGGTAGCGCGGATGTGTCTCATTAACGCGTGGGTGAAGCTTCGCAATGCTGACG 720
Qy 721 GAGATATTTCTCTGCAACAGCAAGGAAATGCGAGGCGGCGGCTGAGGAAAGTACC 780
Db 721 GAGATATTTCTCTGCAACAGCAAGGAAATGCGAGGCGGCGGCTGAGGAAAGTACC 780
Qy 781 GATTCAACAGTGAACACCTGTGAATGTTGCAATACGCGCAATTTGCTGCTACAA 840
Db 781 GATTCAACAGTGAACACCTGTGAATGTTGCAATACGCGCAATTTGCTGCTACAA 840
Qy 841 CGCACGCGAGGTTGCGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
Db 841 CGCACGCGAGGTTGCGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
Qy 901 TTGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
Db 901 TTGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
Qy 961 TTATTCGCGGCAACGATTAATCTGCAATCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
Db 961 TTATTCGCGGCAACGATTAATCTGCAATCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
Qy 1021 ACGCTTCCCGGTCAAGCGGATTAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
Db 1021 ACGCTTCCCGGTCAAGCGGATTAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
Qy 1081 CGTGGGCTTAAGCGATTAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
Db 1081 CGTGGGCTTAAGCGATTAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
Qy 1141 CAGATGCGTGAATAAACGCGCTGTCAATTAATACGCGCGCGGAGAGGATGAACCTGACC 1200
Db 1141 CAGATGCGTGAATAAACGCGCTGTCAATTAATACGCGCGCGGAGAGGATGAACCTGACC 1200
Qy 1201 CTGGCAGGATGTGAAGAGCGAAATGCGGAGGCGATGTGTGTTGCGGAGGTTTAAACGCA 1260
Db 1201 CTGGCAGGATGTGAAGAGCGAAATGCGGAGGCGATGTGTGTTGCGGAGGTTTAAACGCA 1260
Qy 1261 ATCGGATGAAGAGCGATACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1308
Db 1261 ATCGGATGAAGAGCGATACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1308
```

```
RESULT 4
US-11-056-354-1
; Sequence 1, Application US/11056354
; Publication No. US20050246780A1
; GENERAL INFORMATION:
; APPLICANT: SHORT, Jay M.
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin A.
; APPLICANT: BARTON, Nelson R.
; APPLICANT: GARRETT, James B.
; APPLICANT: O'DONOGHUE, Eileen
; APPLICANT: MATHUR, Eric J.
; TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM AND METHODS
; FILE REFERENCE: 56462001803/DI370-9C1
; CURRENT APPLICATION NUMBER: US/11/056,354
; PRIOR APPLICATION NUMBER: US 10/156,660
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1296
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Modified phytase enzyme
US-11-056-354-1

Query Match 99.1%; Score 1296; DB 14; Length 1296;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAGCGATCTTAATCCATTTTATCTCTGATTCGCTTAACCGCGAATTCGCA 60
Db 1 ATGAAAGCGATCTTAATCCATTTTATCTCTGATTCGCTTAACCGCGAATTCGCA 60
Qy 61 TTGCTCAGAGTGAAGCGGAGCTGAAGCTGGAAGTGTGATTTGCACTGCTCATGCT 120
Db 61 TTGCTCAGAGTGAAGCGGAGCTGAAGCTGGAAGTGTGATTTGCACTGCTCATGCT 120
Qy 121 GTGCGTCTCCAAACCAAGGCCGACGAATGATGAGGATGTACCCGACGATGCGCA 180
Db 121 GTGCGTCTCCAAACCAAGGCCGACGAATGATGAGGATGTACCCGACGATGCGCA 180
Qy 181 ACCTGGCGGTTAAACCTGGGTGAGCTGACACCGCGCGGTGTGAGCTTAATCGCTATCTC 240
Db 181 ACCTGGCGGTTAAACCTGGGTGAGCTGACACCGCGCGGTGTGAGCTTAATCGCTATCTC 240
Qy 241 GGAATTACTGCGCTGAGCGCTGCTGAGCGGACGATTTGCTGCTTAATGTGCTGCCG 300
Db 241 GGAATTACTGCGCTGAGCGCTGCTGAGCGGACGATTTGCTGCTTAATGTGCTGCCG 300
Qy 301 CAGTCTGCTGAGTGTGCGATTAATGCTGATGTGCAAGAGCGTACCCGTAACAGCGGAA 360
Db 301 CAGTCTGCTGAGTGTGCGATTAATGCTGATGTGCAAGAGCGTACCCGTAACAGCGGAA 360
Qy 361 GCTTTCGCGCGCGGCTGCGACCTGACCTGCAATTAACCGTAACATCCAGGAGATTCG 420
Db 361 GCTTTCGCGCGCGGCTGCGACCTGACCTGCAATTAACCGTAACATCCAGGAGATTCG 420
```



```
Qy 601 CTTAAAGCTGAGAAACAGAGCAAGAGCTGTTCAATTAAAGCAGGCAATTAACGAACTC 660
Db 788 CTTAAAGCTGAGAAACAGAGCAAGAGCTGTTCAATTAAAGCAGGCAATTAACGAACTC 847
Qy 661 AAGGTGAGCGCGAGCTGTCTCATTTAACCGGTGCGGTAAAGCTTGCATCAATGCTGACG 720
Db 848 AAGGTGAGCGCGAGCTGTCTCATTTAACCGGTGCGGTAAAGCTTGCATCAATGCTGACG 907
Qy 721 GAGATATTTCTCTGCAACAAGCAAGGAAATGCGGAGCGCGGAGTGGGAAAGATCAC 780
Db 908 GAGATATTTCTCTGCAACAAGCAAGGAAATGCGGAGCGCGGAGTGGGAAAGATCAC 967
Qy 781 GATTCAACAAGTGAACACCTTGTGAATTGCAATTAAGCGCAATTTGATTTGCTACAA 840
Db 968 GATTCAACAAGTGAACACCTTGTGAATTGCAATTAAGCGCAATTTGATTTGCTACAA 1027
Qy 841 CGCAGCGCAGAGTGTGCGGAGCGCGGCAACCCCGTTATTAATGATTTGATCAAGACG 900
Db 1028 CGCAGCGCAGAGTGTGCGGAGCGCGGCAACCCCGTTATTAATGATTTGATCAAGACG 1087
Qy 901 TTGAGCGCCGATCCACGCAAAACAGGCGTATGCTGATTAACCACTTCAGTGTG 960
Db 1088 TTGAGCGCCGATCCACGCAAAACAGGCGTATGCTGATTAACCACTTCAGTGTG 1147
Qy 961 TTATTCGCGGACAGATATCTATCTGCAAAATCTGCGGCGGCACTGAGCTCAACTG 1020
Db 1148 TTATTCGCGGACAGATATCTATCTGCAAAATCTGCGGCGGCACTGAGCTCAACTG 1207
Qy 1021 ACGCTTCCCGGTCAAGCGGATTAACAGCGCGGCAAGTGTGAATCTGTGTTGAACGCTG 1080
Db 1208 ACGCTTCCCGGTCAAGCGGATTAACAGCGCGGCAAGTGTGAATCTGTGTTGAACGCTG 1267
Qy 1081 CGTGGGCTAAAGCGATTAACAGCGGATTAAGTTCAGTGTGCTGCTTCAGACCTTTACG 1140
Db 1268 CGTGGGCTAAAGCGATTAACAGCGGATTAAGTTCAGTGTGCTGCTTCAGACCTTTACG 1327
Qy 1141 CAGATCGGTATTAACAGCGGCTGTCAATTAATAGCGCGCGGAGAGGTAAACCTGACC 1200
Db 1328 CAGATCGGTATTAACAGCGGCTGTCAATTAATAGCGCGCGGAGAGGTAAACCTGACC 1387
Qy 1201 CTGGCAGATGTGAAGAGCGAAATGCGGAGGCAATGTGTTGTTGCGAGGTTTAAACGCA 1260
Db 1388 CTGGCAGATGTGAAGAGCGAAATGCGGAGGCAATGTGTTGTTGCGAGGTTTAAACGCA 1447
Qy 1261 ATCGTAATGAAGCAGCATACCGGCGTGCAGTTGAGAT 1300
Db 1448 ATCGTAATGAAGCAGCATACCGGCGTGCAGTTGAT 1487

RESULT 6
US-10-601-319-5
; Sequence 5, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O'Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
; FILE REFERENCE: 09010-029011
; CURRENT APPLICATION NUMBER: US/10/601,319
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
```

```
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 403
; OTHER INFORMATION: n = A,T,C or G
US-10-601-319-5

Query Match 96.9%; Score 1267; DB 8; Length 1901;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 1279; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 1 ATGAAACGATCTTAATCCCATTTTATCTCTCTGATTTCCGTTAAACCCGCAATCTGCA 60
Db 188 ATGAAACGATCTTAATCCCATTTTATCTCTCTGATTTCCGTTAAACCCGCAATCTGCA 247
Qy 61 TTGCTCAGAGTGAAGCCGAGACTGAAGCTGAAAGTGTGATTTGTCAGTCTCATGCT 120
Db 248 TTGCTCAGAGTGAAGCCGAGACTGAAGCTGAAAGTGTGATTTGTCAGTCTCATGCT 307
Qy 121 GTGCTGCTCAACCAAGGCCACGCACTGATGCAAGATGCAACCCAGACGATGGCCA 180
Db 308 GTGCTGCTCAACCAAGGCCACGCACTGATGCAAGATGCAACCCAGACGATGGCCA 367
Qy 181 ACCTGGCGGTTAAACCTGGGTGAGCTGACACCGCGCGTGTGATGCTAAATCGCTATCTC 240
Db 368 ACCTGGCGGTTAAACCTGGGTGAGCTGACACCGCGCGTGTGATGCTAAATCGCTATCTC 427
Qy 241 GGAATTAATCTGGCGTCAAGCTGTGTAAGCGGATTTGCTGCTAAATGTGCTGCCG 300
Db 428 GGAATTAATCTGGCGTCAAGCTGTGTAAGCGGATTTGCTGCTAAATGTGCTGCCG 487
Qy 301 CAGTCTGTGAGTGGGATTAATCTGATGTGACAGAGGTAACCCGTAACAGCGGAA 360
Db 488 CAGTCTGTGAGTGGGATTAATCTGATGTGACAGAGGTAACCCGTAACAGCGGAA 547
Qy 361 GCTTGGCGCGCGGCTGCGACCTGACTGTGCAATTAACCTTAATCCAGGCAAGATACG 420
Db 548 GCTTGGCGCGCGGCTGCGACCTGACTGTGCAATTAACCTTAATCCAGGCAAGATACG 607
Qy 421 TCCAGTCCGATCCGTTATTAATCCCTTAACAACTGGCGTGGCCAACTGATTAACGCG 480
Db 608 TCCAGTCCGATCCGTTATTAATCCCTTAACAACTGGCGTGGCCAACTGATTAACGCG 667
Qy 481 AACGTGCTGACGCGATCTCGAAGGGCAGAGGGTCAATTGCTGACTTTACCGGGCAT 540
Db 668 AACGTGCTGACGCGATCTCGAAGGGCAGAGGGTCAATTGCTGACTTTACCGGGCAT 727
Qy 541 TATCAACCGCGCTTGGCGAATGGAACGGGTCTTAATTTTCCGCAATCAACTGTGTC 600
Db 728 CCGCAACCGCGCTTGGCGAATGGAACGGGTCTTAATTTTCCGCAATCAACTGTGTC 787
Qy 601 CTTAAAGTGAAGAAACAGAGCAAGAGCTGTTCAATTAAAGCAGGCAATTAACATGGAATC 660
Db 788 CTTAAAGTGAAGAAACAGAGCAAGAGCTGTTCAATTAAAGCAGGCAATTAACATGGAATC 847
Qy 661 AAGGTGAGCGCGAGCTGTCTCATTTAACCGGTGCGGTAAAGCTTGCATCAATGCTGACG 720
Db 848 AAGGTGAGCGCGAGCTGTCTCATTTAACCGGTGCGGTAAAGCTTGCATCAATGCTGACG 907
Qy 721 GAGATATTTCTCTGCAACAAGCAAGGAAATGCGGAGCGCGGAGTGGGAAAGATCAC 780
Db 908 GAGATATTTCTCTGCAACAAGCAAGGAAATGCGGAGCGCGGAGTGGGAAAGATCAC 967
```

QY	761	GAATTACACCAAGTGGAAACAACCTTGCTAAGTTTGCATTAACGGCAATTGATTTGGTAAACA	840
Db	968	GATTCAACCACTGGAAACAACCTTGCTAAGTTTGCATTAACGGCAATTGATTTGGTAAACA	10277
QY	841	CGCAGCGCAAGAGTTTGCCCGCAGCGCGCCACCCCGTTATTAGATTGATCAAGACAGCG	900
Db	1028	CGCAGCGCAAGAGTTTGCCCGCAGCGCGCCACCCCGTTATTAGATTGATCAAGACAGCG	10878
QY	901	TTGACGCCCATTCACCGCAAAAACAGCGGTATGGTGTGACATTACCACTTCAGTGTG	960
Db	1088	TTGACGCCCATTCACCGCAAAAACAGCGGTATGGTGTGACATTACCACTTCAGTGTG	11477
QY	961	TTTATCGCGGACACGATACTAATCTGCGAATCTCGCGCGCACCTGAGACTCAACTG	10202
Db	1148	TTTATCGCGGACACGATACTAATCTGCGAATCTCGCGCGCACCTGAGACTCAACTG	12077
QY	1021	AACGCTTCCCGGTCAACGCCGGAATPACACGCCGCCAGGTGTGAACTGTGTGTTGAAACGCTG	10808
Db	1208	AACGCTTCCCGGTCAACGCCGGAATPACACGCCGCCAGGTGTGAACTGTGTGTTGAAACGCTG	12677
QY	1081	CGTCGGCTAAGCGAATPACAGCCACAGTGAATTGAGTTTGCTGAGTCTTCACACTTTACAG	11440
Db	1268	CGTCGGCTAAGCGAATPACAGCCACAGTGAATTGAGTTTGCTGAGTCTTCACACTTTACAG	13277
QY	1141	CAGATGCGTGATTAACCGCGCTGTCAATTAATACGCCGCCCGGAGAGGTGAACCTGACC	12000
Db	1328	CAGATGCGTGATTAACCGCGCTGTCAATTAATACGCCGCCCGGAGAGGTGAACCTGACC	13877
QY	1201	CTGGCAGAGTGTGAAGAGACGAATAATCGCAGAGGCAATGTTGCTTGGCAGGTTTACGGCAA	1260
Db	1388	CTGGCAGAGTGTGAAGAGACGAATAATCGCAGAGGCAATGTTGCTTGGCAGGTTTACGGCAA	14477
QY	1261	ATCGTGAATGAAGCAGCATACCGGCGGTGCAGTTTGAGAT	1300
Db	1448	ATCGTGAATGAAGCAGCATACCGGCGGTGCAGTTTGAGAT	14877

```

1 RESULT 7
2 US-10-933-115-5
3 ; Sequence 5, Application US/10933115
4 ; Publication No. US20050281792N1
5 ; GENERAL INFORMATION:
6 ; APPLICANT: Short, Jay M.
7 ; APPLICANT: Kretz, Keith A.
8 ; APPLICANT: Gray, Kevin A.
9 ; APPLICANT: Barton, Nelson Robert
10 ; APPLICANT: Garrett, James B.
11 ; APPLICANT: O' Donoghue, Eileen
12 ; APPLICANT: Mathur, Eric J.
13 ; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES
14 ; TITLE OF INVENTION: THEREOF
15 ; FILE REFERENCE: 564462001822
16 ; CURRENT APPLICATION NUMBER: US/10/933,115
17 ; CURRENT FILING DATE: 2004-09-01
18 ; PRIOR APPLICATION NUMBER: US/09/866,379
19 ; PRIOR FILING DATE: 2001-05-24
20 ; PRIOR APPLICATION NUMBER: US 09/580,515
21 ; PRIOR FILING DATE: 2000-05-25
22 ; PRIOR APPLICATION NUMBER: US 09/318,528
23 ; PRIOR FILING DATE: 1999-05-25
24 ; PRIOR APPLICATION NUMBER: US 09/291,931
25 ; PRIOR FILING DATE: 1999-04-13
26 ; PRIOR APPLICATION NUMBER: US 09/259,214
27 ; PRIOR FILING DATE: 1999-03-01
28 ; PRIOR APPLICATION NUMBER: US 08/910,798
29 ; PRIOR FILING DATE: 1997-08-13
30 ; NUMBER OF SEQ ID NOS: 10
31 ; SOFTWARE: FastSeq for Windows Version 4.0
32 ; SEQ ID NO 5
33 ; LENGTH: 1901
34 ; TYPE: DNA
35 ; ORGANISM: Escherichia coli
36 ; FEATURE:

```

```

; NAME/KEY: misc_feature
; LOCATION: 403
; OTHER INFORMATION: n = A,T,C or G
US-10-933-115-5

```

Query Match	96.9%	Score 1267;	DB 10;	Length 1901;
Best Local Similarity	98.4%	Pred. No. 0;		
Matches 1279; Conservative	0;	Mismatches 21;	Indels 0;	Gaps 0;

Qy	1	ATGAAAGGAACTTAAATCCCATTTTAACTCTTCGATATCCGTTAAACCCCGAACTCGA	60
Dp	188	ATGAAAGGAACTTAAATCCCATTTTAACTCTTCGATATCCGTTAAACCCCGAACTCGA	247
Qy	61	TTCCGCTCAGATGAGCGGAGCTGAAGCTGGAAGATGGTGGATTTGTCACTGTCATGGT	120
Dp	248	TTCCGCTCAGATGAGCGGAGCTGAAGCTGGAAGATGGTGGATTTGTCACTGTCATGGT	307
Qy	121	GTGCGTGTCCAAACGAGGCCAGCAACTGATGAGATGTCAACCCAGACGATGCGCA	180
Dp	308	GTGCGTGTCCAAACGAGGCCAGCAACTGATGAGATGTCAACCCAGACGATGCGCA	367
Qy	181	ACCTGGCCGGTAAATCTGGGTGACTGTACACCGCGCGGTGGTGGACTTAATCGCTTAATCTC	240
Dp	368	ACCTGGCCGGTAAATCTGGGTGACTGTACACCGCGGNGTGGTGGACTTAATCGCTTAATCTC	427
Qy	241	GGAATTAATGAGCGTCAAGCGTCTGATGAGCGAGCGGATTGTGCGCTAAATGTTGGCTGCCG	300
Dp	428	GGAATTAATGAGCGTCAAGCGTCTGATGAGCGAGCGGATTGTGCGGAAAAAGGGCTGGCCG	487
Qy	301	CAGTCTGGTCAAGTCCGGAATTAATGTCGATGCGACGATACCCGTAAAAACAGCGAA	360
Dp	488	CAGTCTGGTCAAGTCCGGAATTAATGTCGATGCGACGATACCCGTAAAAACAGCGAA	547
Qy	361	GCTTCCGCGCCGGGCTGGCACTGTGTGCAATAACCTGTACATACCGACGAGATACG	420
Dp	548	GCTTCCGCGCCGGGCTGGCACTGTGTGCAATAACCTGTACATACCGACGAGATACG	607
Qy	421	TTCCAGTCCCGATCCGTTATTTAACTCTTAATAAACTGGCGTTTGGCAACTGSAATACGCG	480
Dp	608	TTCCAGTCCCGATCCGTTATTTAACTCTTAATAAACTGGCGTTTGGCAACTGSAATACGCG	667
Qy	481	AACGTGACTGACGCGATCTTCGAGAGGGCAGAGAGGTCAATTGCTGACTTTAACCGGAGAT	540
Dp	668	AACGTGACTGACGCGATCTTCGAGAGGGCAGAGAGGTCAATTGCTGACTTTAACCGGAGAT	727
Qy	541	TATTAACCGCGCTTTCGCGAACTGGAAACGGGTCTTAATTTCCGCAATCAACTTGTGC	600
Dp	728	TATTAACCGCGCTTTCGCGAACTGGAAACGGGTCTTAATTTCCGCAATCAACTTGTGC	787
Qy	601	CTTAAACGTGAGAAACGAGAACGAAGCTGTCAATTAACGAGGCAATTAACATCCGAACTC	660
Dp	788	CTTAAACGTGAGAAACGAGAACGAAGCTGTCAATTAACGAGGCAATTAACATCCGAACTC	847
Qy	661	AAGGTGAGCGCCGACTGTCTCAATTAACCGGTCGGGTAAAGCTTCGATCAATGCTGACG	720
Dp	848	AAGGTGAGCGCCGACTGTCTCAATTAACCGGTCGGGTAAAGCTTCGATCAATGCTGACG	907
Qy	721	GAGATATTTTCTCTGCAACAGCACAGGAAATGCGAGCCGGGCTGGGAGAGATCAC	780
Dp	908	GAGATATTTTCTCTGCAACAGCACAGGAAATGCGAGCCGGGCTGGGAGAGATCAC	967
Qy	781	GATTACACAGATGGAACACTTGTCTAATGTTGATTAACGGCAATTGATTTGCTACAA	840
Dp	968	GATTACACAGATGGAACACTTGTCTAATGTTGATTAACGGCAATTGATTTGCTACAA	1027
Qy	841	CGACGCGCAGAGGTGCGCCGACGCGCGCAACCCGTTATTAAGTTGATCAAGACAGCG	900
Dp	1028	CGACGCGCAGAGGTGCGCGACGCGCGCAACCCGTTATTAAGTTGATCAAGACAGCG	1087
Qy	901	TTGACGCCCATCAACGCAAAAACAGCGGATATGTTGACATTAACCATTTCACTGCTG	960
Dp	1088	TTGACGCCCATCAACGCAAAAACAGCGGATATGTTGACATTAACCATTTCACTGCTG	1147

Qy 961 TTTATGCGCGGACGAGTACTTATCTGGCAAAATCTCGCGGCGGACGACTGGAGCTCACTGG 1020
|
|
|
Db 1148 TTTATGCGCGGACGAGTACTTATCTGGCAAAATCTCGCGGCGGACGACTGGAGCTCACTGG 1207
|
|
|
Qy 1021 ACGCTTCCCGGTGACGCGGATTAACACGCGCGCAGGTGGTGAACGTGTTTGAACGCTGG 1080
|
|
|
Db 1208 ACGCTTCCCGGTGACGCGGATTAACACGCGCGCAGGTGGTGAACGTGTTTGAACGCTGG 1267
|
|
|
Qy 1081 CGTGGCTGAAGCGATTAACAGCACTGATTTTCGCTGCTTCTTCCAACTTTACAG 1140
|
|
|
Db 1268 CGTGGCTGAAGCGATTAACAGCACTGATTTTCGCTGCTTCTTCCAACTTTACAG 1327
|
|
|
Qy 1141 CAGATGCGTGAATAAACGCGCTGTCTTAATAATACGCGCGCGGAGAGGTGAACCTGACC 1200
|
|
|
Db 1328 CAGATGCGTGAATAAACGCGCTGTCTTAATAATACGCGCGCGGAGAGGTGAACCTGACC 1387
|
|
|
Qy 1201 CTGGCAGAGATGTAAGAGCGAAATGCGGAGGCAATGTTGCTTGGCAGATTTTACGCA 1260
|
|
|
Db 1388 CTGGCAGAGATGTAAGAGCGAAATGCGGAGGCAATGTTGCTTGGCAGATTTTACGCA 1447
|
|
|
Qy 1261 ATCGTGAATGAAGCAGCATACCGCGCTGCGAGTTTGAGAT 1300
|
|
|
Db 1448 ATCGTGAATGAAGCAGCATACCGCGCTGCGAGTTTGAT 1487
|
|
|

RESULT 8

US-09-777-566A-1
; Sequence 1, Application US/09777566A
; Patent No. US2001005788A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; TITLE OR INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER370-6
; CURRENT APPLICATION NUMBER: US/09/777, 566A
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: US 09/318, 528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291, 931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259, 214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910, 798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1320)
; NAME/KEY: misc feature
; LOCATION: (1)..(1323)
; OTHER INFORMATION: n is any nucleotide
US-09-777-566A-1

Query Match 96.7%; Score 1265; DB 3; Length 1323;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 1 ATGAAAGCATCTTAATCCCAATTTTATCTCTTGATTCGGTTAACCCCGCAATCTGCA 60
|
|
|
Db 1 ATGAAAGCATCTTAATCCCAATTTTATCTCTTGATTCGGTTAACCCCGCAATCTGCA 60
|
|
|
Qy 61 TTGCTCAGAGTGAAGCGGAGCTGAAGCTGGAAGTGTGATGTTGATGCTGATGCT 120
|
|
|
Db 61 TTGCTCAGAGTGAAGCGGAGCTGAAGCTGGAAGTGTGATGTTGATGCTGATGCT 120
|
|
|
Qy 121 GTGCGTGTCTCAACCAAGCCACGCACTGATGCAAGATGTACCCCAAGACGCAATGCGCA 180
|
|
|

Db 121 GTGCGTGTCTCAACCAAGCCACGCACTGATGCAAGATGTACCCCAAGACGCAATGCGCA 180
|
|
|
Qy 181 AACTGGCCGGTAAACCTGGGTGAGCTGACACCGCGCGGTGGTGAAGCTTAATCGCTATCTC 240
|
|
|
Db 181 AACTGGCCGGTAAACCTGGGTGAGCTGACACCGCGMGTGGTGAAGCTTAATCGCTATCTC 240
|
|
|
Qy 241 GGAATTAATGAGCGTCAAGCTGTGTAGCGGACGGAATTTGCTGCTTAATATGTGCTGCCG 300
|
|
|
Db 241 GGAATTAATCAACCGCAGCGCTGTGTAGCGGACGGAATTTGCTGCGGAAAAAGGCTGCCG 300
|
|
|
Qy 301 CAGTCTGTGAGGTGCGGATTAATTTGCTGATGTGACGAGCGTACCCGTTAAACAGCGAA 360
|
|
|
Db 301 CAGTCTGTGAGGTGCGGATTAATTTGCTGATGTGACGAGCGTACCCGTTAAACAGCGAA 360
|
|
|
Qy 361 GCCTTGCCTCCCGGCTGGGACCTGATGTGGAATTAACCGTACATACCAGCGAGATACG 420
|
|
|
Db 361 GCCTTGCCTCCCGGCTGGGACCTGATGTGGAATTAACCGTACATACCAGCGAGATACG 420
|
|
|
Qy 421 TCCAGTCCCGATCCGTTATTTAATCTCTTAATAAACTGGCTTTGCCAACTGATTAACGCG 480
|
|
|
Db 421 TCCAGTCCCGATCCGTTATTTAATCTCTTAATAAACTGGCTTTGCCAACTGATTAACGCG 480
|
|
|
Qy 481 AACGTGACTGACCGGATCTCTGAGAGGCGAGAGGTCAAATTTGCTGACTTTACCGGCAAT 540
|
|
|
Db 481 AACGTGACTGACCGGATCTCTGAGAGGCGAGAGGTCAAATTTGCTGACTTTACCGGCAAT 540
|
|
|
Qy 541 TATCAACCGCGTTCGCGGAACTGGAAACGGGTGCTTAATTTCCGCAATCAACTGTGCG 600
|
|
|
Db 541 TATCAACCGCGTTCGCGGAACTGGAAACGGGTGCTTAATTTCCGCAATCAACTGTGCG 600
|
|
|
Qy 601 CTTAAACGTGAGAAACAGACGGAAGCTGTCTTAATTAACGAGGCACTTAACATCGGAATC 660
|
|
|
Db 601 CTTAAACGTGAGAAACAGACGGAAGCTGTCTTAATTAACGAGGCACTTAACATCGGAATC 660
|
|
|
Qy 661 AAGGTAGCCCGCACTGTGTCTCAATTAACCGGTGCGGTAAAGCTTCGATCAATGCTGACG 720
|
|
|
Db 661 AAGGTAGCCCGCACTGTGTCTCAATTAACCGGTGCGGTAAAGCTTCGATCAATGCTGACG 720
|
|
|
Qy 721 GAGATATTTCTCTGCAACAAAGCAAGGAATGCGGAGCGGGGTGGGAAGGATACG 780
|
|
|
Db 721 GAGATATTTCTCTGCAACAAAGCAAGGAATGCGGAGCGGGGTGGGAAGGATACG 780
|
|
|
Qy 781 GATTCAACAGAGTGAACACCTTGCTAATGCTTAACCGCAATTTGATTTGCTACAA 840
|
|
|
Db 781 GATTCAACAGAGTGAACACCTTGCTAATGCTTAACCGCAATTTGATTTGCTACAA 840
|
|
|
Qy 841 CGCAGCCGAGAGTGTGCGCGAGCGCGCAACCCCGTTAATTAATTTGATTAACAGACGCG 900
|
|
|
Db 841 CGCAGCCGAGAGTGTGCGCGAGCGCGCAACCCCGTTAATTAATTTGATTAACAGACGCG 900
|
|
|
Qy 901 TTGACGCCCATCCACGCGCAAAAAAGCGGTATGTGTGAATTACCATTCATGAGTCTG 960
|
|
|
Db 901 TTGACGCCCATCCACGCGCAAAAAAGCGGTATGTGTGAATTACCATTCATGAGTCTG 960
|
|
|
Qy 961 TTTATCCCGGACAGTACTTAATCTGCGCAAAATCTCGCGCGGCACTGAGACTCAACTGG 1020
|
|
|
Db 961 TTTATTCGCGGACAGTACTTAATCTGCGCAAAATCTCGCGCGGCACTGAGACTCAACTGG 1020
|
|
|
Qy 1021 ACGCTTCCCGGTGACGCGGATTAACACGCGCGCAGGTGTGAACGTGTTGAACGCTGG 1080
|
|
|
Db 1021 ACGCTTCCCGGTGACGCGGATTAACACGCGCGCAGGTGTGAACGTGTTGAACGCTGG 1080
|
|
|
Qy 1081 CGTGGCTGAAGCGATTAACAGCACTGATTTTCGCTGCTTCTTCCAACTTTACAG 1140
|
|
|
Db 1081 CGTGGCTGAAGCGATTAACAGCACTGATTTTCGCTGCTTCTTCCAACTTTACAG 1140
|
|
|
Qy 1141 CAGATGCGTGAATAAACGCGCTGTCTTAATAATACGCGCGCGGAGAGGTGAACCTGACC 1200
|
|
|
Db 1141 CAGATGCGTGAATAAACGCGCTGTCTTAATAATACGCGCGCGGAGAGGTGAACCTGACC 1200
|
|
|
Qy 1201 CTGGCAGAGTGTGAAGCGAAATGCGCAGGCGCATGTGTTCTGTGGCAGGTTTACGCA 1260
|
|
|
Db 1201 CTGGCAGAGTGTGAAGCGAAATGCGCAGGCGCATGTGTTCTGTGGCAGGTTTACGCA 1260
|
|
|

QY 1261 ATCGGAATGAAGCAGCATACCGGCGTGCAGTTTGAGATCTCATC 1306
DB 1261 ATCGGAATGAAGCAGCATACCGGCGTGCAGTTTGAGATCTCATC 1306

RESULT 9
US-09-866-379-1
Sequence 1, Application US/09866379
Patent No. US20020136754A1
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: SHORT, Jay
APPLICANT: KRETZ, Keith
APPLICANT: GRAY, Kevin
APPLICANT: BARTON, Nelson
APPLICANT: GARRETT, James
APPLICANT: O'DONOGHUE, Eileen
TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
FILE REFERENCE: DIVER1370-7
CURRENT APPLICATION NUMBER: US/09/866,379
CURRENT FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: US 09/580,515
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 09/318,528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: US 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 09/259,214
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: US 08/910,798
PRIOR FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1323
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(1323)
OTHER INFORMATION: n is any nucleotide
NAME/KEY: CDS
LOCATION: (1)..(1323)
OTHER INFORMATION:
US-09-866-379-1

Query Match 96.7%; Score 1265; DB 3; Length 1323;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 ATGAAAGCATTTATCCCATTTTATCTCTTGATTCGGTTAACCCCGCAATCTGCA 60
DB 1 ATGAAAGCATTTATCCCATTTTATCTCTTGATTCGGTTAACCCCGCAATCTGCA 60

QY 61 TTCGCTCAGAGTGAAGCCGGAAGCTGAAAGTGTGTGATTTGTCACTGTCTATGTT 120
DB 61 TTCGCTCAGAGTGAAGCCGGAAGCTGAAAGTGTGTGATTTGTCACTGTCTATGTT 120

QY 121 GTGCGTGTCCAAACCAAGGCGACGCAACTGATGAGAGTCAACCCCGAGAGCATGGCCA 180
DB 121 GTGCGTGTCCAAACCAAGGCGACGCAACTGATGAGAGTCAACCCCGAGAGCATGGCCA 180

QY 181 ACTTGGCCGGTAAACTGGGTGTGCTGACACCGCGAGGTGTGAGCTAATGCGCTATGTC 240
DB 181 ACTTGGCCGGTAAACTGGGTGTGCTGACACCGCGAGGTGTGAGCTAATGCGCTATGTC 240

QY 241 GGAACATTACTGGCGTCAAGGTCTGTGTAGCCGACGATTTCTGCTTAATGTGGTCCCG 300
DB 241 GGAACATTACTCAACGCGCAGGCTGTGTGTAGCCGACGATTTCTGCTTAATGTGGTCCCG 300

QY 301 CAGTCTGAGTCAAGTGCAGTATTTGCTGATGTGACGAGCGTACCCTGTAACAGGCGCA 360
DB 301 CAGTCTGAGTGCAGTATTTGCTGATGTGACGAGCGTACCCTGTAACAGGCGCA 360

DB 301 CAGTCTGAGTGCAGTATTTGCTGATGTGACGAGCGTACCCTGTAACAGGCGCA 360
QY 361 GCGTTGCGCGCGGCGCTGGCACCTGATCTGTGCAATPACCGTACATACCAGCAGATACG 420
DB 361 GCGTTGCGCGCGGCGCTGGCACCTGATCTGTGCAATPACCGTACATACCAGCAGATACG 420

QY 421 TCCAGTCCCGATCCGTTATTTTAATCTCTTAATAAATCGGGGTTTGCAATGATTAACGCG 480
DB 421 TCCAGTCCCGATCCGTTATTTTAATCTCTTAATAAATCGGGGTTTGCAATGATTAACGCG 480

QY 481 AACGTGACTGACGCGATCCTGAGAGGCGAGAGGGTCAATTGCTGACTTTTACCGGCGAT 540
DB 481 AACGTGACTGACGCGATCCTGAGAGGCGAGAGGGTCAATTGCTGACTTTTACCGGCGAT 540

QY 541 TATCAAGCGCGTTCGCGAATCTGAAACGGGTGCTTAATTTTCCGCAATCAACTTTGTC 600
DB 541 CCGCAAGCGCGTTCGCGAATCTGAAACGGGTGCTTAATTTTCCGCAATCAACTTTGTC 600

QY 601 CTTAAGCGAGAAACAGGACGAAGCTGTTCAATTAAGCGAGCATTAACATCGGAATC 660
DB 601 CTTAAGCGAGAAACAGGACGAAGCTGTTCAATTAAGCGAGCATTAACATCGGAATC 660

QY 661 AAGTGAAGCGCGAATCTGTGCTCATTTAACCGGTGCGTAAAGCTGCAATCAATGCTGACG 720
DB 661 AAGTGAAGCGCGAATCTGTGCTCATTTAACCGGTGCGTAAAGCTGCAATCAATGCTGACG 720

QY 721 GAGATATTTCTCTGCAACAGCACAAGGAAATGCCGAGCGGGGTGGGAAAGATCAAC 780
DB 721 GAGATATTTCTCTGCAACAGCACAAGGAAATGCCGAGCGGGGTGGGAAAGATCAAC 780

QY 781 GATTACACCGATGGAACAACCTGTGTAAGTTGCAACGCGCAATTTGATTTGCTACAA 840
DB 781 GATTACACCGATGGAACAACCTGTGTAAGTTGCAACGCGCAATTTGATTTGCTACAA 840

QY 841 CGCAGCGCAGAGTTCGCGCAGCGCGCCACCCCGTATTGATTGATCAAGCAGCG 900
DB 841 CGCAGCGCAGAGTTCGCGCAGCGCGCCACCCCGTATTGATTGATCAAGCAGCG 900

QY 901 TTGACGCCCCATCAACCGCAAAAACAGGCGTATGCTGACATTAACCACTTCACTGCTG 960
DB 901 TTGACGCCCCATCAACCGCAAAAACAGGCGTATGCTGACATTAACCACTTCACTGCTG 960

QY 961 TTGACGCCCCATCAACCGCAAAAACAGGCGTATGCTGACATTAACCACTTCACTGCTG 960
DB 961 TTGACGCCCCATCAACCGCAAAAACAGGCGTATGCTGACATTAACCACTTCACTGCTG 960

QY 961 TTTATTTGCGGACAGATTAATCTGATGCTGCAATTCGCGGCGCACTGAGGCTCAACTGG 1020
DB 961 TTTATTTGCGGACAGATTAATCTGATGCTGCAATTCGCGGCGCACTGAGGCTCAACTGG 1020

QY 1021 ACGCTTCCCGGTCAAGCGGATTAACAGCGCGCGCAAGGTGTAACCTGATTTTGAACGCTGG 1080
DB 1021 ACGCTTCCCGGTCAAGCGGATTAACAGCGCGCGCGCAAGGTGTAACCTGATTTTGAACGCTGG 1080

QY 1081 CGTGGCTTAAGGATTAACAGCGGATTAACAGCGCGCGCGCAAGGTGTAACCTGATTTTGAACGCTGG 1140
DB 1081 CGTGGCTTAAGGATTAACAGCGGATTAACAGCGCGCGCGCGCAAGGTGTAACCTGATTTTGAACGCTGG 1140

QY 1141 CAGATGCGTGAATTAACCGCGCTGTCAATTAATACCGCGCGCGGAGAGGTGAATGTAAC 1200
DB 1141 CAGATGCGTGAATTAACCGCGCTGTCAATTAATACCGCGCGCGGAGAGGTGAATGTAAC 1200

QY 1201 CTGGCAGATGTAAGAGCAGAAATGCGAGGCGATGTGTTGTTGGCAGGTTTACGCA 1260
DB 1201 CTGGCAGATGTAAGAGCAGAAATGCGAGGCGATGTGTTGTTGGCAGGTTTACGCA 1260

QY 1261 ATCGGAATGAAGCAGCATACCGGCGTGCAGTTTGAGATCTCATC 1306
DB 1261 ATCGGAATGAAGCAGCATACCGGCGTGCAGTTTGAGATCTCATC 1306

RESULT 10
US-10-034-985-1
Sequence 1, Application US/10034985
Publication No. US20030049815A1
GENERAL INFORMATION:
APPLICANT: Kretz, Keith

```
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/10/034,985
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US/09/580,515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (1)..(1320)
; NAME/KEY: misc_feature
; LOCATION: (1)..(1323)
; OTHER INFORMATION: n = A,T,C or G
US-10-034-985-1
```

```
Query Match      96.7%; Score 1265; DB 6; Length 1323;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
```

```
QY      1 ATGAAAGCATCTTAATCCATTTTATCTTCGTGATCCGTTAAACCCCGCAATCTGCA 60
DB      1 ATGAAAGCATCTTAATCCATTTTATCTTCGTGATCCGTTAAACCCCGCAATCTGCA 60
QY      61 TTCGTCAGAGTGAAGCCGAGAGCTGAAGCTGGAAGTGTGTGATTTGTCAGTGTCTAGT 120
DB      61 TTCCTCAGAGTGAAGCCGAGAGCTGAAGTGTGTGATTTGTCAGTGTCTAGT 120
QY      121 GTGGGTGCTCAACCAAGGCCAAGCACTGATGACAGATGTCACCCGAGAGCATAGGCA 180
DB      121 GTGGGTGCTCAACCAAGGCCAAGCACTGATGACAGATGTCACCCGAGAGCATAGGCA 180
QY      181 ACCTGCGCGGTAAACTGGTGAAGCTGACACCGCGCGGTGTGAGCTAATGCGCTATCTC 240
DB      181 ACCTGCGCGGTAAACTGGTGAAGCTGACACCGCGCGGTGTGAGCTAATGCGCTATCTC 240
QY      241 GGAATTATCTGCGCTCAAGCTGTGTGAGCCGAGCGGATTTGCTTAATGTGCTGCCCG 300
DB      241 GGAATTATCTGCGCTCAAGCTGTGTGAGCCGAGCGGATTTGCTTAATGTGCTGCCCG 300
QY      301 CAGTCTGGTCAAGTCCGAGTATTTGCTGATGTCGACGAGCGGTACCCGTAACAGGCGAA 360
DB      301 CAGTCTGGTCAAGTCCGAGTATTTGCTGATGTCGACGAGCGGTACCCGTAACAGGCGAA 360
QY      361 GCTTTGCGCGCGGCTGTGCACTGACCTGACCTGCAATTAACCGTACATACCAGGAGATACG 420
DB      361 GCTTTGCGCGCGGCTGTGCACTGACCTGACCTGCAATTAACCGTACATACCAGGAGATACG 420
QY      421 TCACGTCGCCGATCCGTTATTTAATCTCTTAATAAACTGCGCTTTGCAACGTGATTAACGG 480
DB      421 TCACGTCGCCGATCCGTTATTTAATCTCTTAATAAACTGCGCTTTGCAACGTGATTAACGG 480
QY      481 AAGCTGACTGACGAGTCTCTGAGAGGCGAGAGGCTCAATTGCTGATCTTTAACCGGGAT 540
DB      481 AAGCTGACTGACGAGTCTCTGAGAGGCGAGAGGCTCAATTGCTGATCTTTAACCGGGAT 540
QY      541 TATCAAAACGCGGTTTCGGAAGTGGAGCGGGTGTATTTTCGCAATCAAACTTGTGC 600
DB      541 TATCAAAACGCGGTTTCGGAAGTGGAGCGGGTGTATTTTCGCAATCAAACTTGTGC 600
QY      601 CTTAACGCTGAAGAACAGAGCAAGAGCTGTTCAATTAAACGAGGCAATTACATCGGAATC 660
DB      601 CTTAACGCTGAAGAACAGAGCAAGAGCTGTTCAATTAAACGAGGCAATTACATCGGAATC 660
```

```
QY      661 AAGGTGAGCGCGGACCTGTGTCTCATTTAACCGGTGCGGTAAAGCCTCGCATCAATGTGACG 720
DB      661 AAGGTGAGCGCGGACCAATGTCTCATTTAACCGGTGCGGTAAAGCCTCGCATCAATGTGACG 720
QY      721 GAGATATTTCTCTGCAACAAGCAAGAGGATGCGGAGCGGAGTGGGAAAGGATCAC 780
DB      721 GAGATATTTCTCTGCAACAAGCAAGAGGATGCGGAGCGGAGTGGGAAAGGATCAC 780
QY      781 GATTCAACCAAGTGAACACCTTGTCTAAGTTTGATTAACCGCAATTTGATTGCTCA 840
DB      781 GATTCAACCAAGTGAACACCTTGTCTAAGTTTGATTAACCGCAATTTGATTGCTCA 840
QY      841 CGCAGCGCAAGGTTGCGCGGAGCGGCGGACCCCGTTATTTAATTGATCAAGACGCG 900
DB      841 CGCAGCGCAAGGTTGCGCGGAGCGGCGGACCCCGTTATTTGATTTGATCAAGACGCG 900
QY      901 TTGACGCCCATTCACCGCAAAAAACAGCGGTATGTGACATTTACCCACTTCAGTGTG 960
DB      901 TTGACGCCCATTCACCGCAAAAAACAGCGGTATGTGACATTTACCCACTTCAGTGTG 960
QY      961 TTTATGCGCGGACACCATTAATCTGTGCAAAATCTCGCGCGGCACTGAGACTCACTGG 1020
DB      961 TTTATGCGCGGACACCATTAATCTGTGCAAAATCTCGCGCGGCACTGAGACTCACTGG 1020
QY      1021 ACGCTTCCGGTGAAGCGGATTAACAGCGCGGAGGTGTGAACCTGTTTGAACGCTGG 1080
DB      1021 ACGCTTCCGGTGAAGCGGATTAACAGCGCGGAGGTGTGAACCTGTTTGAACGCTGG 1080
QY      1081 CGTGGCTGAAGCGATTAACAGCGGAGTTCAGGTTTCGTGTCTTCCAGACTTTACAG 1140
DB      1081 CGTGGCTGAAGCGATTAACAGCGGAGTTCAGGTTTCGTGTCTTCCAGACTTTACAG 1140
QY      1141 CAGATGCTGATTAACAGCGGCTGTCTTAATTAACCGCGCGGAGAGGTGAACCTGACC 1200
DB      1141 CAGATGCTGATTAACAGCGGCTGTCTTAATTAACCGCGCGGAGAGGTGAACCTGACC 1200
QY      1201 CTGGCAGATGTGAAGCGGAAATGCGCAGGCGATGTTCTGTTGCGAGGTTTACGCA 1260
DB      1201 CTGGCAGATGTGAAGCGGAAATGCGCAGGCGATGTTCTGTTGCGAGGTTTACGCA 1260
QY      1261 ATCGTGAATGAAGCAGCATACCGGCGTGCAGTTTGATGATCTCATC 1306
DB      1261 ATCGTGAATGAAGCAGCATACCGGCGTGCAGTTTGATGATCTCATC 1306

RESULT 11
US-10-430-356-1
; Sequence 1, Application US/10430356
; Publication No. US20030232041A1
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/10/430,356
; CURRENT FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: US/09/580,515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
```

```
/ NAME/KEY: CDS
/ LOCATION: (1)...(1320)
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1323)
/ OTHER INFORMATION: n = A,T,C or G
US-10-430-356-1
```

```
Query Match      96.7%; Score 1265; DB 7; Length 1323;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
```

```
QY 1 ATGAAAGCATCTTAATCCATTTTATCTCTTCTGATTCCTGTAACCCGCAATGTGA 60
DB 1 ATGAAAGCATCTTAATCCATTTTATCTCTTCTGATTCCTGTAACCCGCAATGTGA 60
QY 61 TTGCTCAGAGTGAAGCCGGAAGCTGAAAGCTGAAAGTGTGTGATTTGTCAATGCT 120
DB 61 TTGCTCAGAGTGAAGCCGGAAGCTGAAAGTGTGTGATTTGTCAATGCTGATGCT 120
QY 121 GGGCGGTCTCCAAACCAAGGCCAAGCAATGATGCAAGAGTCAACCCCAAGCCATG 180
DB 121 GGGCGGTCTCCAAACCAAGGCCAAGCAATGATGCAAGAGTCAACCCCAAGCCATG 180
QY 181 ACCTGCGCGGTAAACCTGAGGTGAGCTGACACCGCGCGGTGTGAGCTAATGCGCT 240
DB 181 ACCTGCGCGGTAAACCTGAGGTGAGCTGACACCGCGCGGTGTGAGCTAATGCGCT 240
QY 241 GGAATTAATCTGAGCTGACGCTGTGATGAGCCGAAGATTTGCTGCTAAATGTG 300
DB 241 GGAATTAATCTGAGCTGACGCTGTGATGAGCCGAAGATTTGCTGCTAAATGTG 300
QY 301 CAGTGTGATGAGTGTGCAATTTTGTGATGATGCAAGAGTCAACCCGTAAACAGCC 360
DB 301 CAGTGTGATGAGTGTGCAATTTTGTGATGATGCAAGAGTCAACCCGTAAACAGCC 360
QY 361 GCTTGTGCGCGCGGTGAGCACTGATGATGCAATTAACCGTCAATACCAAGCAAT 420
DB 361 GCTTGTGCGCGCGGTGAGCACTGATGATGCAATTAACCGTCAATACCAAGCAAT 420
QY 421 TCCAGTCCGATCGGTATTTAATCTCTTAAACCTGCGGTGTGCAATGATTAAC 480
DB 421 TCCAGTCCGATCGGTATTTAATCTCTTAAACCTGCGGTGTGCAATGATTAAC 480
QY 481 AACGTGATGACGCAATCTCTGAGAGGCGCAAGAGGTCAATGCTGACTTTACCG 540
DB 481 AACGTGATGACGCAATCTCTGAGAGGCGCAAGAGGTCAATGCTGACTTTACCG 540
QY 541 TATCAAAACGCGGTTCGGAACCTGGAACGCGGTCTTAATTTTCCGCAATCAAT 600
DB 541 TATCAAAACGCGGTTCGGAACCTGGAACGCGGTCTTAATTTTCCGCAATCAAT 600
QY 601 CTTAAACGTGAAGAACAGAGCAAGAGTCTTCAATTAACGAGGCAATTAACAT 660
DB 601 CTTAAACGTGAAGAACAGAGCAAGAGTCTTCAATTAACGAGGCAATTAACAT 660
QY 661 AAGGTGAGCGCGCAATGTCTCATTAACCGGTGCGGTGAAGTGTGCAATGCT 720
DB 661 AAGGTGAGCGCGCAATGTCTCATTAACCGGTGCGGTGAAGTGTGCAATGCT 720
QY 721 GAGATATTTCTCTGCAACAAAGCAAGGGAATGCGGAGCGCGGTGTGAGGAAG 780
DB 721 GAGATATTTCTCTGCAACAAAGCAAGGGAATGCGGAGCGCGGTGTGAGGAAG 780
QY 781 GATTCAACAGTGAAGCACTTGTCTAAGTTGCAATTAACGCGCAATTTTATTT 840
DB 781 GATTCAACAGTGAAGCACTTGTCTAAGTTGCAATTAACGCGCAATTTTATTT 840
QY 841 CGCAGCGCAAGAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
DB 841 CGCAGCGCAAGAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
QY 901 TTGACGCCCATTCACCGCAAAAACAGGCGATGTGTGATTCACCACTTCAATG 960
```

```
DB 901 TTGACGCCCATTCACCGCAAAAACAGGCGATGTGTGATTCACCACTTCAATG 960
QY 961 TTTATGCGCGGACACGATATCTAATCTGCAAAATCTCGCGGCGCACTGAGCT 1020
DB 961 TTTATGCGCGGACACGATATCTAATCTGCAAAATCTCGCGGCGCACTGAGCT 1020
QY 1021 AAGCTTCCCGGTCAACCGGATTAACACCGCGCAAGTGTGAACTGCTTTTGA 1080
DB 1021 AAGCTTCCCGGTCAACCGGATTAACACCGCGCAAGTGTGAACTGCTTTTGA 1080
QY 1081 CGTCCGCTTAAGCGATTAACAGCGGATGATTTGAGTTTGTGCTGCTTTCAG 1140
DB 1081 CGTCCGCTTAAGCGATTAACAGCGGATGATTTGAGTTTGTGCTGCTTTCAG 1140
QY 1141 CAGATGCGGTATTAACCGCGGTGTCTAATTAATACCGCGCGGAGAGGTGA 1200
DB 1141 CAGATGCGGTATTAACCGCGGTGTCTAATTAATACCGCGCGGAGAGGTGA 1200
QY 1201 CTGCGAGATGTGAAGAGCGAAATGCGAGGCGCATGTCTTGTGCGAGTTT 1260
DB 1201 CTGCGAGATGTGAAGAGCGAAATGCGAGGCGCATGTCTTGTGCGAGTTT 1260
QY 1261 ATCGTAATGAAGACAGCATACCGCGGTGAGTTGAGATCTATC 1306
DB 1261 ATCGTAATGAAGACAGCATACCGCGGTGAGTTGAGATCTATC 1306
```

```
RESULT 12
US-10-601-319-1
/ Sequence 1, Application US/10601319
/ Publication No. US20040091968A1
/ GENERAL INFORMATION:
/ APPLICANT: Short, Jay M.
/ APPLICANT: Kretz, Keith A.
/ APPLICANT: Gray, Kevin A.
/ APPLICANT: Barton, Nelson Robert
/ APPLICANT: Garrett, James B.
/ APPLICANT: O' Donoghue, Eileen
/ APPLICANT: Mathur, Eric J.
/ TITLE OF INVENTION: RECOMBINANT PHRYASES AND METHODS OF MAKING
/ FILE REFERENCE: 09010-029011
/ CURRENT APPLICATION NUMBER: US/10/601,319
/ PRIOR FILING DATE: 2003-06-20
/ PRIOR APPLICATION NUMBER: US 09/866,379
/ PRIOR FILING DATE: 2001-05-24
/ PRIOR APPLICATION NUMBER: US 09/580,515
/ PRIOR FILING DATE: 2000-05-25
/ PRIOR APPLICATION NUMBER: US 09/318,528
/ PRIOR FILING DATE: 1999-05-25
/ PRIOR APPLICATION NUMBER: US 09/291,931
/ PRIOR FILING DATE: 1999-04-13
/ PRIOR APPLICATION NUMBER: US 09/259,214
/ PRIOR FILING DATE: 1999-03-01
/ PRIOR APPLICATION NUMBER: US 08/910,798
/ PRIOR FILING DATE: 1997-08-13
/ NUMBER OF SEQ ID NOS: 10
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1
/ LENGTH: 1323
/ TYPE: DNA
/ ORGANISM: Escherichia coli
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)...(1320)
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 216
/ OTHER INFORMATION: n = A,T,C or G
US-10-601-319-1
Query Match      96.7%; Score 1265; DB 8; Length 1323;
```

Best Local Similarity 98.0%; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

```

Qy 1 ATGAAAGGATCTTAATCCCAATTTTATCTCTTGATTCCTGTTAAACCCCGCATCTGCA 60
Db 1 ATGAAAGGATCTTAATCCCAATTTTATCTCTTGATTCCTGTTAAACCCCGCATCTGCA 60
Qy 61 TTCGCTCAGAGTGAGCCGAGCTGAGAGTGAGAGTGAGTGTGATTCAGTCTGATAGT 120
Db 61 TTCGCTCAGAGTGAGCCGAGCTGAGAGTGAGAGTGAGTGTGATTCAGTCTGATAGT 120
Qy 121 GTGCGTCTCCAAACCAAGGCCACGCACTGATGCAAGATGTCACCCCGAGACGATGCGCA 180
Db 121 GTGCGTCTCCAAACCAAGGCCACGCACTGATGCAAGATGTCACCCCGAGACGATGCGCA 180
Qy 181 AACTGGCCGGTAAACCTGGGTGAGCTGACACCGCGCGTGGTGGAGCTTAATGCCCTATCTC 240
Db 181 AACTGGCCGGTAAACCTGGGTGAGCTGACACCGCGCGTGGTGGAGCTTAATGCCCTATCTC 240
Qy 241 GGAACATTAATGAGCGCTGAGCTGAGTCCGACGAGATGCTGCTTAAATGTTGCTGCCG 300
Db 241 GGAACATTAATGAGCGCTGAGCTGAGTCCGACGAGATGCTGCTTAAATGTTGCTGCCG 300
Qy 301 CAGTCTGCTCAGTCTGAGTATGCTGATGCTGACGAGCGTAAACCCGTTAAACAGGCGAA 360
Db 301 CAGTCTGCTCAGTCTGAGTATGCTGATGCTGACGAGCGTAAACCCGTTAAACAGGCGAA 360
Qy 361 GCGTTCCGCGCGCGGCTGGACCTGACCTGCAATTAACCGTACATPACCCAGGAGATACG 420
Db 361 GCGTTCCGCGCGCGGCTGGACCTGACCTGCAATTAACCGTACATPACCCAGGAGATACG 420
Qy 421 TCCAGTCCCGATCCGTTATTTATCTCTTAAAACTGCGCTTGGCACTGATTAACGCG 480
Db 421 TCCAGTCCCGATCCGTTATTTATCTCTTAAAACTGCGCTTGGCACTGATTAACGCG 480
Qy 481 AACGTGACTGACGCGATCTTGAAGAGGCGACGAGGCTCAATTTGCTGACTTTAACGGGAT 540
Db 481 AACGTGACTGACGCGATCTTGAAGAGGCGACGAGGCTCAATTTGCTGACTTTAACGGGAT 540
Qy 541 TATCAAAACGGGCTTCCGGAAGTGGAGGCTTATTTTCCGCAATCAAACTTTGGC 600
Db 541 TATCAAAACGGGCTTCCGGAAGTGGAGGCTTATTTTCCGCAATCAAACTTTGGC 600
Qy 601 CTTAAACGTGAGAAACAGAGCAAAAGCTGTCATTTAACGACGACTTACATCCGAACTC 660
Db 601 CTTAAACGTGAGAAACAGAGCAAAAGCTGTCATTTAACGACGACTTACATCCGAACTC 660
Qy 661 AAGGTGAGCGCGACTGCTCTCATTTAACGGGTGCTGAGCTTCGATCAATGCTGACG 720
Db 661 AAGGTGAGCGCGACTGCTCTCATTTAACGGGTGCTGAGCTTCGATCAATGCTGACG 720
Qy 721 GAGATATTTCTCCGGAACAAGCAAGGAAATGCCGAGCGCGGTGGGAAAGATCC 780
Db 721 GAGATATTTCTCCGGAACAAGCAAGGAAATGCCGAGCGCGGTGGGAAAGATCC 780
Qy 781 GATTCAACACAGTGAACACCTGTCTAATTGCTAATACGCGCAATTTGATTTGCTACAA 840
Db 781 GATTCAACACAGTGAACACCTGTCTAATTGCTAATACGCGCAATTTGATTTGCTACAA 840
Qy 841 CGCAGCGCAAGGTTGCTGCGAGCGCGCGCAACCCCGTTATTGATTTGATCAAGACGCG 900
Db 841 CGCAGCGCAAGGTTGCTGCGAGCGCGCGCAACCCCGTTATTGATTTGATCAAGACGCG 900
Qy 901 TTGACGCGCCATCAACGCAAAACAGGCGTATGTTGATTAATCAATCCACTTCAGTGTG 960
Db 901 TTGACGCGCCATCAACGCAAAACAGGCGTATGTTGATTAATCAATCCACTTCAGTGTG 960
Qy 961 TTATTCGCGGACACGATATCTGCAATCTCGCGCGGACAGTGGAGCTCAACTGCG 1020
Db 961 TTATTCGCGGACACGATATCTGCAATCTCGCGCGGACAGTGGAGCTCAACTGCG 1020
Qy 1021 AGCGTTCCGGTCAAGCGGATTAACGCGCGCAGGTGTTGAACTGCTGTTGAACTGCG 1080
Db 1021 AGCGTTCCGGTCAAGCGGATTAACGCGCGCAGGTGTTGAACTGCTGTTGAACTGCG 1080

```

```

Db 1021 ACCTTCCTCGGTCAGCCGATTAACACCGCCGACAGTGTGTAATGATGTTGAACTGCTG 1080
Qy 1081 CGTCGCTAAGCGATTAACAGCCAGTGAATTCAGGTTTCGCTGCTTCCAGACTTTACAG 1140
Db 1081 CGTCGCTAAGCGATTAACAGCCAGTGAATTCAGGTTTCGCTGCTTCCAGACTTTACAG 1140
Qy 1141 CAGATGCTGATTAACACCGCGCTGCTCATTAATACCGCGCCGAGAGGTGAAACTGAC 1200
Db 1141 CAGATGCTGATTAACACCGCGCTGCTCATTAATACCGCGCCGAGAGGTGAAACTGAC 1200
Qy 1201 CTGGCAGATGTGAAGCGCAATTCGCGAGGCAATGCTGCTGCTGCTGCTGCTGCTGCTG 1260
Db 1201 CTGGCAGATGTGAAGCGCAATTCGCGAGGCAATGCTGCTGCTGCTGCTGCTGCTGCTG 1260
Qy 1261 ATCGTAATGAAGCAGCATACCGCGCTGAGTTGAGTCTCATC 1306
Db 1261 ATCGTAATGAAGCAGCATACCGCGCTGAGTTGAGTCTCATC 1306

```

RESULT 13

US-10-933-115-1

Sequence 1, Application US/10933115

Publication No. US20050281792A1

GENERAL INFORMATION:

APPLICANT: Short, Jay M.

APPLICANT: Kretz, Keith A.

APPLICANT: Gray, Kevin A.

APPLICANT: Barton, Nelson Robert

APPLICANT: O' Donoghue, Eileen

APPLICANT: Mathur, Eric J.

TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES

FILE REFERENCE: 564462001822

CURRENT APPLICATION NUMBER: US/10/933,115

CURRENT FILING DATE: 2004-09-01

PRIOR APPLICATION NUMBER: US/09/866,379

PRIOR FILING DATE: 2001-05-24

PRIOR APPLICATION NUMBER: US 09/580,515

PRIOR FILING DATE: 2000-05-25

PRIOR APPLICATION NUMBER: US 09/318,528

PRIOR FILING DATE: 1999-05-25

PRIOR APPLICATION NUMBER: US 09/291,931

PRIOR FILING DATE: 1999-04-13

PRIOR APPLICATION NUMBER: US 09/259,214

PRIOR FILING DATE: 1999-03-01

PRIOR APPLICATION NUMBER: US 08/910,798

PRIOR FILING DATE: 1997-08-13

NUMBER OF SEQ ID NOS: 10

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1

LENGTH: 1323

TYPE: DNA

ORGANISM: Escherichia coli

FEATURE:

NAME/KEY: CDS

LOCATION: (1)...(1320)

FEATURE:

NAME/KEY: misc_feature

LOCATION: 216

OTHER INFORMATION: n = A,T,C or G

US-10-933-115-1

Query Match 96.7%; Score 1265; DB 10; Length 1323;

Best Local Similarity 98.0%; Pred. No. 0;

Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

```

Qy 1 ATGAAAGGATCTTAATCCCAATTTTATCTCTTGATTCCTGTTAAACCCCGCATCTGCA 60
Db 1 ATGAAAGGATCTTAATCCCAATTTTATCTCTTGATTCCTGTTAAACCCCGCATCTGCA 60
Qy 61 TTCGCTCAGAGTGAGCCGAGCTGAGAGTGAGAGTGAGTGTGATTCAGTCTGATAGT 120
Db 61 TTCGCTCAGAGTGAGCCGAGCTGAGAGTGAGAGTGAGTGTGATTCAGTCTGATAGT 120

```

Db 61 TTGCTCAGAGTGAGCCGAGCTGAAGCTGAAAGTGTGATGTCAGTGCATGGT 120
Qy 121 GTGGGCTCCAAAGCCGCACTGATGCAAGATGTCACCCCAAGACGATGGCA 180
Db 121 GTGGGCTCCAAAGCCGCACTGATGCAAGATGTCACCCCAAGACGATGGCA 180
Qy 181 ACCCTGCGGCTGAACTGGGTGAGTGAACAACGCGGCTGATGAGCTTAATCGCTATCTC 240
Db 181 ACCCTGCGGCTGAACTGGGTGAGTGAACAACGCGGCTGATGAGCTTAATCGCTATCTC 240
Qy 241 GGAATTAATGCGCTGACGCTGTGTGATGCGACGATGCTGCTTAATGTGCTGCGCG 300
Db 241 GGAATTAATGCGCTGACGCTGTGTGATGCGACGATGCTGCTTAATGTGCTGCGCG 300
Qy 301 CAGTCTGCTGAGTGTGCGGATTAATGCTGATGTCAGACGCTAACCCGTAAACAGCGCGAA 360
Db 301 CAGTCTGCTGAGTGTGCGGATTAATGCTGATGTCAGACGCTAACCCGTAAACAGCGCGAA 360
Qy 361 GCTTTCGCGCGGCTGACCTGACTGTGCAATACGCTAATCCAGGAGATAG 420
Db 361 GCTTTCGCGCGGCTGACCTGACTGTGCAATACGCTAATCCAGGAGATAG 420
Qy 421 TCCAGTCCCGATCCGTTATTTAATCTCTAATAAAGCTGCTTGCACATGATTAACGCG 480
Db 421 TCCAGTCCCGATCCGTTATTTAATCTCTAATAAAGCTGCTTGCACATGATTAACGCG 480
Qy 481 AACGTGATGAGCGGCTCTCTGAGAGGCGAGAGGCTCAATTGCTGATCTTTAACCGGCGAT 540
Db 481 AACGTGATGAGCGGCTCTCTGAGAGGCGAGAGGCTCAATTGCTGATCTTTAACCGGCGAT 540
Qy 541 TATCAAAAGCGGCTTTCGCGAATGGAAGGCTTAATTTTCCGCAATCAAACTTGTGC 600
Db 541 TATCAAAAGCGGCTTTCGCGAATGGAAGGCTTAATTTTCCGCAATCAAACTTGTGC 600
Qy 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGACGACATTAACATCGAATCTC 660
Db 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGACGACATTAACATCGAATCTC 660
Qy 661 AAGGTGACCGGCTGCTGTCTCAATTAACCGGTGCGGTAAAGCTGCTGATTAATGCTGACG 720
Db 661 AAGGTGACCGGCTGCTGTCTCAATTAACCGGTGCGGTAAAGCTGCTGATTAATGCTGACG 720
Qy 721 GAGATATTTCTCTCAACAGACAGGGAATGCGGAGCGGCGGAGGAGGATCAACC 780
Db 721 GAGATATTTCTCTCAACAGACAGGGAATGCGGAGCGGCGGAGGAGGATCAACC 780
Qy 781 GATTCAACCAAGTGAACAACCTTGTGAATTTGCAATACGCGCAATTTTATTTGCTACAA 840
Db 781 GATTCAACCAAGTGAACAACCTTGTGAATTTGCAATACGCGCAATTTTATTTGCTACAA 840
Qy 841 CGCAGCGCAAGGTTGCGCGACGCGCGCACCCGTTATGATTTGATTCATGCAAGG 900
Db 841 CGCAGCGCAAGGTTGCGCGACGCGCGCACCCGTTATGATTTGATTCATGCAAGG 900
Qy 901 TTGAGCGCCCAATCCACCGGAAACAGAGGCTATGATGATTAATGATTCAGACAGG 960
Db 901 TTGAGCGCCCAATCCACCGGAAACAGAGGCTATGATGATTAATGATTCAGACAGG 960
Qy 961 TTTATGCGCGGACAGATATTAATCTGCAATCTCGCGCGGCACTGAGGCTCAACTGG 1020
Db 961 TTTATGCGCGGACAGATATTAATCTGCAATCTCGCGCGGCACTGAGGCTCAACTGG 1020
Qy 1021 ACCGTTCCCGGTCAAGCGGATTAACGCGCGCAAGGTGATTAATGATTTGAAAGCTGG 1080
Db 1021 ACCGTTCCCGGTCAAGCGGATTAACGCGCGCAAGGTGATTAATGATTTGAAAGCTGG 1080
Qy 1081 CGTGGCTTAAGGATTAACGAGGATTAATGATTTGCTGCTTTCAGCTTTTAAAG 1140
Db 1081 CGTGGCTTAAGGATTAACGAGGATTAATGATTTGCTGCTTTCAGCTTTTAAAG 1140
Qy 1141 CAGATGCTGATTAACGCGGCTGTCTATTAATACGCGCGGAGAGGATGAACCTGACC 1200
Db 1141 CAGATGCTGATTAACGCGGCTGTCTATTAATACGCGCGGAGAGGATGAACCTGACC 1200

Qy 1201 CTGGCAGATGTTGAAAGGCAATGCGCAGGCAATGTGTTGCTGCGAGCTTTTACGAA 1260
Db 1201 CTGGCAGATGTTGAAAGGCAATGCGCAGGCAATGTGTTGCTGCGAGCTTTTACGAA 1260
Qy 1261 ATCGTAATGAAAGCAGCATACCGGCGTGCAGTTTGATGATCTCATC 1306
Db 1261 ATCGTAATGAAAGCAGCATACCGGCGTGCAGTTTGATGATCTCATC 1306

RESULT 14
US-10-282-122A-7167
; Sequence 7167, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Foreyn, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 7167
; LENGTH: 1299
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-282-122A-7167

Query Match 96.6%; Score 1264; DB 8; Length 1299;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 1276; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 1 ATGAAAGGATCTTAATCCATTTTATCTCTGATTCGTTAAACCGCAATCTCA 60
Db 1 ATGAAAGGATCTTAATCCATTTTATCTCTGATTCGTTAAACCGCAATCTCA 60
Qy 61 TTGCTCAGAGTGAGCCGAGCTGAAGCTGGAAGTGTGATGTCAGTGTCTATGCT 120
Db 61 TTGCTCAGAGTGAGCCGAGCTGAAGCTGGAAGTGTGATGTCAGTGTCTATGCT 120
Qy 121 GTGGTGTCTCAACCAAGGCCACGCACTGATGCAAGATGTCACCCCAAGACGATGGCA 180

```
Db 121 GTGGTGTCTCCAAACCAAGCCACGCACTGATGAGGATGTCACCCCGAAGCATGGCCA 180
Qy 181 ACCTGGCCGGTAAACTGGGTGAGTGCACCCGCGGTGGTGAATGAGCTATCTC 240
Db 181 ACCTGGCCGGTAAACTGGGTGAGTGCACCCGCGGTGGTGAATGAGCTATCTC 240
Qy 241 GGACATTACTGGCGGTCAAGCTGTGGTAGCCGACGGAATTGCTGCTTAAATGATGGCTGCCG 300
Db 241 GGACATTACTGGCGGTCAAGCTGTGGTAGCCGACGGAATTGCTGCTTAAATGATGGCTGCCG 300
Qy 301 CAGTCTGGTCAAGTTCGCGATTAATGCTGATGTGACGAGCGTAAACCGTAAACAGGCCAA 360
Db 301 CAGTCTGGTCAAGTTCGCGATTAATGCTGATGTGACGAGCGTAAACCGTAAACAGGCCAA 360
Qy 361 GCCTTGGCCGCGCGGTGGCACTGACCTGTGCAATACCGTACATACCAGCGATACG 420
Db 361 GCCTTGGCCGCGCGGTGGCACTGACCTGTGCAATACCGTACATACCAGCGATACG 420
Qy 421 TCCAGTCCCGATCCGTTATTAATCTCTAAACCTGGCGTTTGGCAACTGGATPACGCG 480
Db 421 TCCAGTCCCGATCCGTTATTAATCTCTAAACCTGGCGTTTGGCAACTGGATPACGCG 480
Qy 481 AACGTGACTGACGCGATCTCTGAGAGGGCAGAGGGGTCAATGCTGACTTTAACGGGCAT 540
Db 481 AACGTGACTGACGCGATCTCTGAGAGGGCAGAGGGGTCAATGCTGACTTTAACGGGCAT 540
Qy 541 TATCAAAAGCGGTTTCCGGAACCTGGAACGGGTGCTTAATTTTCCGCACTCAACTTGGC 600
Db 541 TATCAAAAGCGGTTTCCGGAACCTGGAACGGGTGCTTAATTTTCCGCACTCAACTTGGC 600
Qy 601 CTTAAACGTGAGAAACAGAGCAAAAGCTGTCATTAAACGACGAGCATTAACATCCGAACTC 660
Db 601 CTTAAACGTGAGAAACAGAGCAAAAGCTGTCATTAAACGACGAGCATTAACATCCGAACTC 660
Qy 661 AAGGTGAGCGCGCATGTGTCTCAATTAACCGGTGCGTAAAGCTTCGATCAATGCTGACG 720
Db 661 AAGGTGAGCGCGCATGTGTCTCAATTAACCGGTGCGTAAAGCTTCGATCAATGCTGACG 720
Qy 721 GAGATATTTTCCCTGCAACAGCAAGGAAATGCCGAGCCGGGTGGGAAAGATCAACC 780
Db 721 GAGATATTTTCCCTGCAACAGCAAGGAAATGCCGAGCCGGGTGGGAAAGATCAACC 780
Qy 781 GATTCAACACAGTGAACACCTGTGAATTTGCAACGCGCAATTTGATTTGCTACAA 840
Db 781 GATTCAACACAGTGAACACCTGTGAATTTGCAACGCGCAATTTGATTTGCTACAA 840
Qy 841 CGCAGCGCAAGGTTGCCGAGCCGCGCACCCCGTTATTGATTTGATCAAGACAGCG 900
Db 841 CGCAGCGCAAGGTTGCCGAGCCGCGCACCCCGTTATTGATTTGATCAAGACAGCG 900
Qy 901 TTGAGCGCCCATCAACCGCAAAACAGGCGATGATGATTAACCACTTCAGTGTG 960
Db 901 TTGAGCGCCCATCAACCGCAAAACAGGCGATGATGATTAACCACTTCAGTGTG 960
Qy 961 TTATATCGCGGACAGATTAATCTGGCAATCTCGCGGCGCACTGAGCTCAACTG 1020
Db 961 TTATATCGCGGACAGATTAATCTGGCAATCTCGCGGCGCACTGAGCTCAACTG 1020
Qy 1021 ACGCTTCCCGGTCAAGCGGATPACAGCGCGCAAGGTGTGAATGTGTTTGAACGCTGG 1080
Db 1021 ACGCTTCCCGGTCAAGCGGATPACAGCGCGCAAGGTGTGAATGTGTTTGAACGCTGG 1080
Qy 1081 CGTGGCGTAAACGATTAACAGCAAGGATTAAGGTTTGGCTGCTTCAACATTTACG 1140
Db 1081 CGTGGCGTAAACGATTAACAGCAAGGATTAAGGTTTGGCTGCTTCAACATTTACG 1140
Qy 1141 CAGATGCGTATTAACAGCGCGCTGTCAATTAATACGCGCGGAGAGGTGAACCTGACC 1200
Db 1141 CAGATGCGTATTAACAGCGCGCTGTCAATTAATACGCGCGGAGAGGTGAACCTGACC 1200
Qy 1201 CTGCGAGATGTGAAGACGAAATGCGAGGCGATGTGTTGTTGACGAGTTTACGCA 1260
Db 1201 CTGCGAGATGTGAAGACGAAATGCGAGGCGATGTGTTGTTGACGAGTTTACGCA 1260
```

```
Db 1201 CTGCGAGATGTGAAGACGAAATGCGAGGCGATGTGTTGTTGACGAGTTTACGCA 1260
Qy 1261 ATCGTAATGAAGACCATTAACCGCGCGTGAAGTTG 1296
Db 1261 ATCGTAATGAAGACCATTAACCGCGCGTGAAGTTG 1296

RESULT 15
US-09-866-379-6
; Sequence No, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER370-7
; CURRENT FILING DATE: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-6

Query Match 96.6%; Score 1263.8; DB 3; Length 1901;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 1277; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 1 ATGAAGGAGATCTTAATCCCATTTTATCTTCTGATTCGGTTAACCCCGCAATCTGCA 60
Db 188 ATGAAGGAGATCTTAATCCCATTTTATCTTCTGATTCGGTTAACCCCGCAATCTGCA 247
Qy 61 TTGCTCAGAGTGAAGCCGAGCTGAAGCTGGAAGTGTGATTTGCTCACTGCTCATGT 120
Db 248 TTGCTCAGAGTGAAGCCGAGCTGAAGCTGGAAGTGTGATTTGCTCACTGCTCATGT 307
Qy 121 GTGGTGTCTCCAAACCAAGGCCAGCACTGATGAGATGTCAACCCAGACGCAATGGCCA 180
Db 121 GTGGTGTCTCCAAACCAAGGCCAGCACTGATGAGATGTCAACCCAGACGCAATGGCCA 180
Qy 308 GTGGTGTCTCCAAACCAAGGCCAGCACTGATGAGATGTCAACCCAGACGCAATGGCCA 367
Db 308 GTGGTGTCTCCAAACCAAGGCCAGCACTGATGAGATGTCAACCCAGACGCAATGGCCA 367
Qy 181 ACCTGGCCGGTAAACTGGGTGAGTGCACCCGCGGTGGTGAATGAGCTATCTC 240
Db 181 ACCTGGCCGGTAAACTGGGTGAGTGCACCCGCGGTGGTGAATGAGCTATCTC 240
Qy 368 ACGTGGCCGGTAAACTGGGTGAGTGCACCCGCGGTGGTGAATGAGCTATCTC 427
Db 368 ACGTGGCCGGTAAACTGGGTGAGTGCACCCGCGGTGGTGAATGAGCTATCTC 427
Qy 241 GGACATTACTGGCGGTCAAGCTGTGGTAGCCGACGGAATTGCTGCTTAAATGATGGCTGCCG 300
Db 241 GGACATTACTGGCGGTCAAGCTGTGGTAGCCGACGGAATTGCTGCTTAAATGATGGCTGCCG 300
Qy 428 GGACATTACTGGCGGTCAAGCTGTGGTAGCCGACGGAATTGCTGCTTAAATGATGGCTGCCG 487
Db 428 GGACATTACTGGCGGTCAAGCTGTGGTAGCCGACGGAATTGCTGCTTAAATGATGGCTGCCG 487
Qy 301 CAGTCTGGTCAAGTTCGCGATTAATGCTGATGTGACGAGCGTAAACCGTAAACAGGCCAA 360
Db 301 CAGTCTGGTCAAGTTCGCGATTAATGCTGATGTGACGAGCGTAAACCGTAAACAGGCCAA 360
Qy 488 CAGTCTGGTCAAGTTCGCGATTAATGCTGATGTGACGAGCGTAAACCGTAAACAGGCCAA 547
Db 488 CAGTCTGGTCAAGTTCGCGATTAATGCTGATGTGACGAGCGTAAACCGTAAACAGGCCAA 547
```

QY 361 GCCTTCGCGCCGCGGCTGGCACTGACTGTCATTAACCGTACATTAACCGAGGATACG 420
DB 548 GCCTTCGCGCGCGGCTGGCACTGACTGTCATTAACCGTACATTAACCGAGGATACG 607
QY 421 TCCAGTCCCGATCCGTTATTTAAATCTTAATAAAGCTGCTTGGCACTGATTAACGCG 480
DB 608 TCCAGTCCCGATCCGTTATTTAAATCTTAATAAAGCTGCTTGGCACTGATTAACGCG 667
QY 481 AACGTGACTGACCGCATCTCTGAGAGGCGAGAGGCTCAATTGCTGACTTTACCGGCAAT 540
DB 668 AACGTGACTGACCGCATCTCTGAGAGGCGAGAGGCTCAATTGCTGACTTTACCGGCAAT 727
QY 541 TATCAAAACGCGCTTTGCGCAACTGAAACGCGGCTTAATTTTCCGCAATCAACTGTGC 600
DB 728 CCGCAAAACGCGCTTTGCGCAACTGAAACGCGGCTTAATTTTCCGCAATCAACTGTGC 787
QY 601 CTTAAACGTGAGAAACAGGACGAAACCTGTTCAATTAACGAGGCAATTAACATCGAACTC 660
DB 788 CTTAAACGTGAGAAACAGGACGAAACCTGTTCAATTAACGAGGCAATTAACATCGAACTC 847
QY 661 AAGGTGAGCGCGCACTGTCTCTATTAAACGCGTGGCTAAAGCTCGCATCAATGCTGACG 720
DB 848 AAGGTGAGCGCGCACTGTCTCTATTAAACGCGTGGCTAAAGCTCGCATCAATGCTGACG 907
QY 721 GAGATATTTCTCTGCAACAGGACAGGAAATGCGGAGCGGCGGAGGAGGATCACG 780
DB 908 GAGATATTTCTCTGCAACAGGACAGGAAATGCGGAGCGGCGGAGGAGGATCACG 967
QY 781 GATTCAACACAGTGAACAACCTTGTAGTTTGCAATTAACGCGCAATTTGATTGCTACAA 840
DB 968 GATTCAACACAGTGAACAACCTTGTAGTTTGCAATTAACGCGCAATTTGATTGCTACAA 1027
QY 841 CGCAGCGCGAGAGTTGCCCGCAGCGCGCAACCCGTTATTAGATTGATCAAGACAGCG 900
DB 1028 CGCAGCGCGAGAGTTGCCCGCAGCGCGCAACCCGTTATTAGATTGATCAAGACAGCG 1087
QY 901 TTGACGCCCCCATTCACCGCAAAACAGGCGTATGAGTGAACATTAACCACTTCAAGTCG 960
DB 1088 TTGACGCCCCCATTCACCGCAAAACAGGCGTATGAGTGAACATTAACCACTTCAAGTCG 1147
QY 961 TTTATCGCGGACACGATATCTATCTGCAAAATCTCGCGCGGCACTGAGCTCAACTG 1020
DB 1148 TTTATCGCGGACACGATATCTATCTGCAAAATCTCGCGCGGCACTGAGCTCAACTG 1207
QY 1021 ACGCTTCCCGGTCAAGCCGATTAACAGCGCGCAAGGTGTGAACGTGTGTTGAACGCTG 1080
DB 1208 ACGCTTCCCGGTCAAGCCGATTAACAGCGCGCAAGGTGTGAACGTGTGTTGAACGCTG 1267
QY 1081 CGTCGCGTAAAGGATTAACAGCGAGTGAATTCAGGTTTGGCTGCTTCCAGACTTACAG 1140
DB 1268 CGTCGCGTAAAGGATTAACAGCGAGTGAATTCAGGTTTGGCTGCTTCCAGACTTACAG 1327
QY 1141 CAGATGCGTATTAACAGCGCGTGTCTATTAATAACGCGCGCGAGAGGTGAACCTGACC 1200
DB 1328 CAGATGCGTATTAACAGCGCGTGTCTATTAATAACGCGCGCGAGAGGTGAACCTGACC 1387
QY 1201 CTGCGAGATGTGAAGAGGAAATGCGCAGGSCATGTGTTGTTGCGAGGTTTAAAGCA 1260
DB 1388 CTGCGAGATGTGAAGAGGAAATGCGCAGGSCATGTGTTGTTGCGAGGTTTAAAGCA 1447
QY 1261 ATCGTGAATGAAGACGCACTACCGGCGTGCATTTGAGAT 1300
DB 1448 ATCGTGAATGAAGACGCACTACCGGCGTGCATTTGAT 1487

Search completed: June 14, 2006, 12:41:32
Job time : 159.84 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2006, 10:00:16 ; Search time 889.897 Seconds

(without alignments)
10248.037 Million cell updates/sec

Title: US-10-601-319-9

Perfect score: 1308
Sequence: 1 atgaagcagcattcaatccc.....gcatttgagatcctaccta 1308

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

N_Geneseq.8:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001s:*
- 5: geneseqn2002s:*
- 6: geneseqn2003s:*
- 7: geneseqn2004s:*
- 8: geneseqn2005s:*
- 9: geneseqn2006s:*
- 10: geneseqn2007s:*
- 11: geneseqn2008s:*
- 12: geneseqn2009s:*
- 13: geneseqn2010s:*
- 14: geneseqn2011s:*
- 15: geneseqn2012s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1308	100.0	1308	10	ADc87742 DNA encod
2	1308	100.0	1308	12	ADd50299 Escherich
3	1308	100.0	1308	15	Aee75420 E. coli B
4	1296	99.1	1296	14	Aed50817 Escherich
5	1294.2	98.9	1294	15	Aee75422 E. coli B
6	1268.6	97.0	1268	12	ADd50296 Escherich
7	1267	96.9	1267	12	ADd50295 Escherich
8	1267	96.9	1267	12	Aee75416 E. coli B
9	1265	96.7	1265	5	Aac88885 Escherich
10	1265	96.7	1265	6	ADd36473 Escherich
11	1265	96.7	1265	6	ADd25460 Escherich
12	1265	96.7	1265	12	ADd19445 E. coli B
13	1265	96.7	1265	12	ADd50291 Escherich
14	1265	96.7	1265	12	ADd50301 Kangaroo
15	1265	96.7	1265	12	ADd50303 Escherich
16	1265	96.7	1265	15	Aee75412 E. coli B
17	1264	96.6	1264	3	ABK12514 DNA encod
18	1264	96.6	1264	8	ACA19297 Prokaryot

19	1264	96.6	1264	12	ADd16137 Escherich
20	1263.8	96.6	1263	6	ADd25463 Escherich
21	1263.8	96.6	1263	8	ADd19449 E. coli K
22	1263.8	96.6	1263	10	ADc87744 DNA encod
23	1263.8	96.6	1263	12	ADd50297 Escherich
24	1263.8	96.6	1263	14	ADd50819 Escherich
25	1263.8	96.6	1263	15	Aee75418 E. coli B
26	1263.8	96.6	1263	15	Aee75417 E. coli B
27	1260.8	96.4	1260	12	ADd16139 Shigella
28	1260	96.3	1304	14	Aed50824 Escherich
29	1259.2	96.3	3470	3	Aac68298 R15/APPA
30	1259.2	96.3	4060	3	Aac68297 R15/APPA
31	1259.2	96.3	6116	3	Aac68297 R15/APPA
32	1259.2	96.3	6708	3	Aac68295 R15/APPA
33	1259.2	96.3	17732	3	Aac68300 Lama2/APP
34	1259.2	96.3	20623	3	Aac68294 Lama2/APP
35	1258.4	96.2	1315	10	ADd102197 DNA encod
36	1251	95.6	1901	8	ADd19452 E. coli K
37	1249.6	95.5	1299	15	Aee75418 Escherich
38	1246.2	95.3	5421	3	Aac68299 R15/APPA
39	1243.2	95.0	1296	12	ADd16138 Shigella
40	1243.2	95.0	1299	13	ADd76362 Phytase A
41	1233.6	94.3	1486	4	Aad6831 E. coli a
42	1228	93.9	1489	3	Aaa28216 E. coli a
43	1228	93.9	1489	8	ACC57672 Escherich
44	1222.4	93.5	1486	4	Aad6832 E. coli a
45	1222.4	93.5	1486	8	ACC57673 Escherich

ALIGNMENTS

RESULT 1	ADc87742	ADc87742 standard; DNA; 1308 BP.
ID	ADc87742	
XX	ADc87742	
AC	ADc87742	
XX	ADc87742	
DT	01-JAN-2004	(first entry)
XX	01-JAN-2004	
DS	DNA encoding the modified Escherichia coli phytase.	
XX	Phytase; food supplement; enzyme delivery matrix; soybean meal;	
KW	thermotolerance; thermostability; kernel; phytate;	
KW	myo-inositol-hexaphosphate; inositol; inorganic phosphate;	
KW	thermotolerant; feed value; digestion; gene; de.	
XX	Synthetic.	
OS	Escherichia coli.	
XX	Escherichia coli.	
XX	Key	Location/Qualifiers
FT	CDS	1..1308
FT		/*tag= a
FT		/product= "Modified phytase"
XX	US2003103958-A1.	
XX	05-JUN-2003.	
PD	24-MAY-2002; 2002US-00156660.	
XX	13-AUG-1997; 97US-00910798.	
PR	01-MAR-1999; 99US-00259214.	
PR	13-APR-1999; 99US-00291931.	
PR	25-MAY-1999; 99US-00318528.	
PR	25-MAY-2000; 2000US-00580515.	
PR	24-MAY-2001; 2001US-00866379.	
XX	(DIVE-) DIVERSA CORP.	
PA	Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue B;	
PI	Mathur EJ,	
PI		
XX		

DR WPI: 2003-787039/74.
 DR P-PSDB; ADC87743.
 XX
 PT New nucleic acid encoding a polypeptide having phytase activity, useful
 PT in improving the feeding value of phytate rich ingredients or as an aid
 PT in phytate digestion.
 PS
 PS Claim 9; SEQ ID NO 1; 113pp; English.
 XX
 XX The invention discloses a new isolated or recombinant nucleic acid which
 CC encodes a polypeptide having a phytase activity. Also claimed is a
 CC nucleic acid probe, an amplification primer sequence pair, an expression
 CC cassette comprising the nucleic acid, a vector comprising the nucleic
 CC acid, a transgenic non-human animal or plant, or its seed, comprising the
 CC nucleic acid, an antisense oligonucleotide, inhibiting the translation of
 CC a phytase message in a cell, a heterodimer comprising the polypeptide and
 CC a second domain, an array comprising immobilised polypeptide or nucleic
 CC acid, a hybridoma comprising an antibody that specifically binds to the
 CC polypeptide, a food supplement for an animal, an edible enzyme delivery
 CC matrix, an edible pellet comprising a granule edible carrier and the
 CC polypeptide, a feed composition, a soybean meal, isolating or identifying
 CC the polypeptide, making an anti-phytase antibody, producing a recombinant
 CC polypeptide, determining whether a compound binds to the polypeptide,
 CC identifying a modulator, whole cell engineering of new or modified
 CC phenotypes by using real-time metabolic flux analysis, increasing
 CC the resistance or the thermostability of the phytase polypeptide, increasing
 CC the resistance of the phytase polypeptide to enzymatic inactivation in a
 CC digestive system and processing of corn and sorghum kernels. The phytase
 CC activity comprises catalysts of phytate (myo-inositol-hexaphosphate) to
 CC inositol and inorganic phosphate and is thermostolerant. The nucleic acid
 CC is useful in improving the feeding value of phytate rich ingredients or
 CC as an aid in phytate digestion. The sequence presented is the DNA
 CC encoding the modified *Escherichia coli* phytase.
 CC
 XX
 SQ Sequence 1308 BP; 315 A; 345 C; 357 G; 291 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1308; DB 10; Length 1308;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 481 AACGTGACTGACCGGATCCTCGAGAGGCGAGAGGGTCAATTGTGTGACTTTACCGGGCAT 540
 Qy 541 TATTAACGGCGTTTCCGGAACCTGGAACGGGTCTTAATTTTCCGCAATCAAACTGTGC 600
 Db 541 TATCAACGGCGTTTCCGGAACCTGGAACGGGTCTTAATTTTCCGCAATCAAACTGTGC 600
 Qy 601 CTTAAAGTGAGAAACAGGACGAAGCTGTTCAATTACGAGGATTAACATCGGAATC 660
 Db 601 CTTAAAGTGAGAAACAGGACGAAGCTGTTCAATTACGAGGATTAACATCGGAATC 660
 Qy 661 AAGGTGAGCGCGCATGTGTCTCAATTAAACGGTGGTAAAGCTTCGATCAATGCTACG 720
 Db 661 AAGGTGAGCGCGCATGTGTCTCAATTAAACGGTGGTAAAGCTTCGATCAATGCTACG 720
 Qy 721 GAGATATTTTCTCTCGCAACAAGCAGAGGAATGCCGAGCCGGGTGGGAAGATAC 780
 Db 721 GAGATATTTTCTCTCGCAACAAGCAGAGGAATGCCGAGCCGGGTGGGAAGATAC 780
 Qy 781 GATTCAACCAAGTGAACACCTTGCTAAGTTTGATACGCGCAATTGATTTGCTACAA 840
 Db 781 GATTCAACCAAGTGAACACCTTGCTAAGTTTGATACGCGCAATTGATTTGCTACAA 840
 Qy 841 CGCAGCGCAGAGGTTGCGCGCAGCGCGCAACCCGTTATTAATTGATCAAGACAGCG 900
 Db 841 CGCAGCGCAGAGGTTGCGCGCAGCGCGCAACCCGTTATTAATTGATCAAGACAGCG 900
 Qy 901 TTGACGCCCATCCACCGCAAAAACAGGCGTATGTGTGACATTAACCATTCAGTCTG 960
 Db 901 TTGACGCCCATCCACCGCAAAAACAGGCGTATGTGTGACATTAACCATTCAGTCTG 960
 Qy 961 TTTATGCCCGGACAGATTAATCTGGCAAAATCTCGCGCGCGCATGAGCTCAACTGG 1020
 Db 961 TTTATGCCCGGACAGATTAATCTGGCAAAATCTCGCGCGCGCATGAGCTCAACTGG 1020
 Qy 1021 ACGCTTCGGGTGACCGCGATTAACACGCGCGCAGGTGTGAATGCTGTTGAACGCTGG 1080
 Db 1021 ACGCTTCGGGTGACCGCGATTAACACGCGCGCAGGTGTGAATGCTGTTGAACGCTGG 1080
 Qy 1081 CGTGGCTAAGGCAATACACCGAGTATTCAGGTTTCGTGTCTCCGACATTTACAG 1140
 Db 1081 CGTGGCTAAGGCAATACACCGAGTATTCAGGTTTCGTGTCTCCGACATTTACAG 1140
 Qy 1141 CAGATGGGTATTAACGCGCGTGTCAATTAATACGCGCGCGGAGAGGTGAATGACC 1200
 Db 1141 CAGATGGGTATTAACGCGCGTGTCAATTAATACGCGCGCGGAGAGGTGAATGACC 1200
 Qy 1201 CTGGCAGATGTGAAGAGCGAATGCGCAGGCGATGTGTTGTTGCGAGGTTTACGCAA 1260
 Db 1201 CTGGCAGATGTGAAGAGCGAATGCGCAGGCGATGTGTTGTTGCGAGGTTTACGCAA 1260
 Qy 1261 ATGTGAATGAAGACGATACCGGCGTGTGAGTTGAGATCTACTCA 1308
 Db 1261 ATGTGAATGAAGACGATACCGGCGTGTGAGTTGAGATCTACTCA 1308

RESULT 2
 ID ADO50299 standard; DNA; 1308 BP.
 XX
 AC ADO50299;
 DT 29-JUL-2004 (first entry)
 XX
 DE *Escherichia coli* K12 *appa* phytase mutant 819PH59 DNA.
 XX
 KW *Appa* phytase; bacteria; thermal tolerance; protease stability; foodstuff;
 FM phytate; animal feed; fish feed; dough; baking; gene; mutant; de.
 XX
 OS *Escherichia coli*; K12.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers

```

CDS
1..1296
/*tag= a
/product= "Mutant phytase"
/note= "Does not include stop codon"
/partial
replace(202,T)
/*tag= b
/note= "Corresponds to position 389 in the wild type
sequence"
replace(203,G)
/*tag= c
/note= "Corresponds to position 390 in the wild type
sequence"
replace(250,C)
/*tag= d
/note= "Corresponds to position 437 in the wild type
sequence"
replace(251,A)
/*tag= e
/note= "Corresponds to position 438 in the wild type
sequence"
replace(252,A)
/*tag= f
/note= "Corresponds to position 439 in the wild type
sequence"
replace(255,C)
/*tag= g
/note= "Corresponds to position 442 in the wild type
sequence"
replace(283,G)
/*tag= h
/note= "Corresponds to position 470 in the wild type
sequence"
replace(285,G)
/*tag= i
/note= "Corresponds to position 472 in the wild type
sequence"
replace(289,G)
/*tag= j
/note= "Corresponds to position 476 in the wild type
sequence"
replace(290,A)
/*tag= k
/note= "Corresponds to position 477 in the wild type
sequence"
replace(291,G)
/*tag= l
/note= "Corresponds to position 478 in the wild type
sequence"
replace(502,A)
/*tag= m
/note= "Corresponds to position 689 in the wild type
sequence"
replace(503,G)
/*tag= n
/note= "Corresponds to position 690 in the wild type
sequence"
replace(504,C)
/*tag= o
/note= "Corresponds to position 691 in the wild type
sequence"
replace(541,C)
/*tag= p
/note= "Corresponds to position 728 in the wild type
sequence"
replace(542,G)
/*tag= q
/note= "Corresponds to position 729 in the wild type
sequence"
replace(543,G)
/*tag= r
/note= "Corresponds to position 730 in the wild type
sequence"

```

```

FT mutation replace(676,A)
FT /*tag= g
FT /note= "Corresponds to position 863 in the wild type
FT sequence"
FT mutation replace(677,A)
FT /*tag= t
FT /note= "Corresponds to position 864 in the wild type
FT sequence"
FT mutation replace(829,T)
FT /*tag= u
FT /note= "Corresponds to position 1016 in the wild type
FT sequence"
XX
XX US2004091968-A1.
XX
XX 13-MAY-2004.
XX
XX 20-JUN-2003; 2003US-00601319.
XX
XX 13-AUG-1997; 97US-00910798.
XX 01-MAR-1999; 99US-00259214.
XX 13-APR-1999; 99US-00291931.
XX 25-MAY-1999; 99US-00318528.
XX 25-MAY-2000; 2000US-00580515.
XX 24-MAY-2001; 2001US-00866379.
XX
XX (SHOR/) SHORT J M.
XX (KREB/) KREBZ K.
XX (GRAY/) GRAY K A.
XX (BART/) BARTON N R.
XX (GAR/) GARRETT J B.
XX (ODON/) O'DONOGHUE E.
XX (MATH/) MATHER E J.
XX
XX Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E,
XX Mather EJ;
XX
XX MPI. 2004-374952/35.
XX
XX P-PSDB; AD050300.
XX
XX Producing phytase, involves providing nucleic acid derived from bacteria
XX encoding polypeptide having phytase activity, and expressing nucleic acid
XX in yeast.
XX
XX Claim 3; SEQ ID NO 9; 74pp; English.
XX
XX The invention relates to a method for producing phytase that involves
XX providing a nucleic acid encoding phytase derived from a bacteria, and
XX expressing the nucleic acid in a yeast under conditions that allow
XX expression of the enzyme in the yeast. The invention also relates to
XX modified phytase enzyme which has improved thermal tolerance and protease
XX stability at low pH. The phytase enzyme can be used in foodstuffs to
XX improve the feeding value of phytate rich ingredients, and in diet of
XX numerous animals including mammals, fowls and fishes, commercially
XX significant mammals such as pigs, goats, laboratory rodents, commercially
XX significant avian species such as chicken, ducks, doves, parrot, etc.,
XX commercially farmed fish such as guppy, zebrafish, molly, swordtail,
XX etc., in dough making and baking, in dietary aids for animals. The method
XX provides easy manufacture of the active ingredient loaded biocompatible
XX composition, higher yields and loading efficiency. The phytase
XX incorporated in to the dietary aid is safe for animals. The present
XX sequence is Bescherichia coli K12 apra phytase mutant DNA.
XX
XX Sequence 1308 BP; 315 A; 345 C; 357 G; 291 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 1308; DB 12; Length 1308;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 ATGAAGGAGATCTTAATCCATTCTTCCTGATTCGTTAACCCCGCAATCTGCA 60
XX |||
XX 1 ATGAAGGAGATCTTAATCCATTCTTCCTGATTCGTTAACCCCGCAATCTGCA 60
XX
XX Db

```

```

Qy 61 TTCCCTGAGAGTGGCCGAGCTGAGCTGGAAGTGTGATGATGTCATGTCATGATGAT 120
Db 61 TTGCTCAGAGTGAAGCCGAGCTGAGTGAAGTGTGATGATGATGATGATGATGATGATGAT 120
Qy 121 GTGGGTGTCTCAACCAAGGCCACGCACTGATGACGAGATGTCACCCGACGATGGCCA 180
Db 121 GTGGGTGTCTCAACCAAGGCCACGCACTGATGACGAGATGTCACCCGACGATGGCCA 180
Qy 181 ACCTGGCCGGTAAACTGGGTGAGTGAACACCGCGCGGTGGTGAAGCTAATGCGCTATCTC 240
Db 181 ACCTGGCCGGTAAACTGGGTGAGTGAACACCGCGCGGTGGTGAAGCTAATGCGCTATCTC 240
Qy 241 GGACATTACTGCGCTGACGCTGCTGATGACCGACGATGCTGCTGCTGCTGCTGCTGCTG 300
Db 241 GGACATTACTGCGCTGACGCTGCTGATGACCGACGATGCTGCTGCTGCTGCTGCTGCTG 300
Qy 301 CAGTCTGTGTCAGGTGCGGATTTATGCTGATGTCGACGAGCGTACCCTGTAACACGCGCA 360
Db 301 CAGTCTGTGTCAGGTGCGGATTTATGCTGATGTCGACGAGCGTACCCTGTAACACGCGCA 360
Qy 361 GCCTTCGCGCGCGGGCTGGACCTGACCTGTGCAATTAACCGTACATACCAGGACGATACG 420
Db 361 GCCTTCGCGCGCGGGCTGGACCTGACCTGTGCAATTAACCGTACATACCAGGACGATACG 420
Qy 421 TCCAGTCCCGATCCGTTATTTATCTCTTAAACCTGCGCTTGGCCACTGATTAACGCG 480
Db 421 TCCAGTCCCGATCCGTTATTTATCTCTTAAACCTGCGCTTGGCCACTGATTAACGCG 480
Qy 481 AACGTGACTGACCGGATCTTGAAGAGGCGAGAGGGGTCAATTTGCTGACTTTACCGGGCAT 540
Db 481 AACGTGACTGACCGGATCTTGAAGAGGCGAGAGGGGTCAATTTGCTGACTTTACCGGGCAT 540
Qy 541 TATGAAACGGGCTTTGCGGAACTGGAACGGGCTGTTATTTTCCGCAATCAACCTGTCG 600
Db 541 TATGAAACGGGCTTTGCGGAACTGGAACGGGCTGTTATTTTCCGCAATCAACCTGTCG 600
Qy 601 CTTAAACGTGAGAAACAGAGCAAAAGCTTTCAATTAAACGACGCAATTACATCCGAACTC 660
Db 601 CTTAAACGTGAGAAACAGAGCAAAAGCTTTCAATTAAACGACGCAATTACATCCGAACTC 660
Qy 661 AAGGTGAGCGCGGACTGTGTCTCATTTAACCGGTGCGGTGAACCTTGCATCAATGCTGACG 720
Db 661 AAGGTGAGCGCGGACTGTGTCTCATTTAACCGGTGCGGTGAACCTTGCATCAATGCTGACG 720
Qy 721 GAGATATTTCTCTGCAACAGACACAGGGAATGCGGAGCGCGGGTGGGAAAGATCACG 780
Db 721 GAGATATTTCTCTGCAACAGACACAGGGAATGCGGAGCGCGGGTGGGAAAGATCACG 780
Qy 781 GATTCAACACAGTGAACACCTTGTGTAAGTTGCAATAACGCGCAATTTGATTTGCTACAA 840
Db 781 GATTCAACACAGTGAACACCTTGTGTAAGTTGCAATAACGCGCAATTTGATTTGCTACAA 840
Qy 841 CGCACGCCAGAGGTGGCCCGGACGCGCGCACCCCGTTATTAAGATTGATTAACAGACGCG 900
Db 841 CGCACGCCAGAGGTGGCCCGGACGCGCGCACCCCGTTATTAAGATTGATTAACAGACGCG 900
Qy 901 TTGACGCCCATTCACCGCAAAAACAGCGGTATGTTGATTAACCACTTCAGTGTGCTG 960
Db 901 TTGACGCCCATTCACCGCAAAAACAGCGGTATGTTGATTAACCACTTCAGTGTGCTG 960
Qy 961 TTTAATCGCCGACACGATTAATCTGTGCAATCTCGCGCGGCGCATCTGAGCTCAACTGG 1020
Db 961 TTTAATCGCCGACACGATTAATCTGTGCAATCTCGCGCGGCGCATCTGAGCTCAACTGG 1020
Qy 1021 ACGCTTCCCGGTGACGCGGATTAACAGCGCGGCGAGGTGTGAACCTGTGTGGAACGCTGG 1080
Db 1021 ACGCTTCCCGGTGACGCGGATTAACAGCGCGGCGAGGTGTGAACCTGTGTGGAACGCTGG 1080
Qy 1081 CGTGGGCTTAAGCGATTAACAGCGCGGATTAACAGCGCGGATTAACAGCGCGGATTAACAG 1140
Db 1081 CGTGGGCTTAAGCGATTAACAGCGCGGATTAACAGCGCGGATTAACAGCGCGGATTAACAG 1140
Qy 1141 CAGATGCGTGAATAAACGCGCGCTGTCAATTAATACGCGCGCGCGGAGAGGTGAACCTGACC 1200

```

```

Db 1141 CAGATGCGTGAATAAACGCGCTGTCAATTAATACGCGCGCGGAGAGGTGAACCTGACC 1200
Qy 1201 CTGGCAGAGATGTGAAGGCCAAATGCGCAGAGGCAATGTTCTGTTGGCAGGTTTAACCAA 1260
Db 1201 CTGGCAGAGATGTGAAGGCCAAATGCGCAGAGGCAATGTTCTGTTGGCAGGTTTAACCAA 1260
Qy 1261 ATCGTAATGAAGCAGCATACCGGCGTGCAGATTGAGATCTCATCTTA 1308
Db 1261 ATCGTAATGAAGCAGCATACCGGCGTGCAGATTGAGATCTCATCTTA 1308

RESULT 3
AEE75420
ID AEE75420 standard; DNA; 1308 BP.
XX
AC AEE75420;
XX
DT 23-FEB-2006 (first entry)
XX
DE E. coli B modified appa phytase 819PH59 DNA sequence SEQ ID NO: 9.
XX
KW ds; coding sequence; mutant; thermostable; phytase;
KW protein stabilization; pharmaceutical; fertilizer; cramp; cell culture;
KW osteoporosis; osteopathic; food; appa.
XX
OS Escherichia coli K12.
XX
SY Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..1296
FT FT /*tag= a
FT FT /product= "modified phytase appa"
XX
PN US2005281792-A1.
XX
PD 22-DEC-2005.
XX
PF 01-SEP-2004; 2004US-00933115.
XX
PR 13-AUG-1997; 97US-00910798.
PR 01-MAR-1999; 99US-00259214.
PR 13-APR-1999; 99US-00291931.
PR 25-MAY-1999; 99US-00318528.
PR 25-MAY-2000; 2000US-00580515.
PR 24-MAY-2001; 2001US-00866379.
XX
PA (SHOR/) SHORT J M.
PA (KRET/) KRETZ K A.
PA (GRAY/) GRAY K A.
PA (BART/) BARTON N R.
PA (GARR/) GARRETT J B.
PA (ODON/) ODOGHOUE E.
PA (BAUM/) BAUM W.
PA (ROBE/) ROBERTSON D E.
PA (ZORN/) ZORNER P.
XX
PI Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, Odonoghue E;
PI Baum W, Robertson DE, Zorner P;
XX
XX WPI; 2006-055022/06.
DR P-PSDB; AEE75421.
XX
PT Formulation useful as dietary supplement for treating, preventing or
PT reversing osteoporosis and bone loss, and preventing muscle cramps,
PT comprises one or more polypeptides having phytase activity.
XX
PS Example 2; SEQ ID NO 9; 82bp; English.
XX
CC The present sequence is that of a polynucleotide encoding a novel
CC modified phytase derived from the Escherichia coli appa protein, with
CC improved thermal tolerance and protease stability compared to the wild-
CC type. The present invention relates to a novel formulation, comprising at

```

CC least one polypeptide having phytase activity, where the polypeptide is a
 CC modified sequence (AEB75421) derived from the *Escherichia coli* K-12 *apra*
 CC phytase (AEB75419). The modification of the enzyme, comprising a series
 CC of mutations, improves the thermal tolerance and protease stability of
 CC the protein. The specification also claims a pharmaceutical composition;
 CC a kit; an immobilized phytase; a fertilizer or soil additive; a liquid
 CC supplement for preventing muscle cramps; a hydrating agent; a tissue
 CC culture or cell culture media; and a plant food additive, all comprising
 CC the novel phytase, and a method of reducing pollution and increasing
 CC nutrient availability in an environment or environmental sample by
 CC degrading environmental phytic acid. The novel enzyme is a dietary
 CC supplement useful for treating, preventing or reversing osteoporosis or
 CC bone loss, and preventing muscle cramps. The liquid supplement is useful
 CC for preventing muscle cramps. The method is useful for reducing pollution
 CC and increasing nutrient availability in an environment or environmental
 CC sample by degrading environmental phytic acid, where the environment or
 CC environmental sample comprises a soil or a body of water. The immobilized
 CC phytase is useful in foodstuffs for improving the feeding value of
 CC phytate rich ingredients.

CC Sequence 1308 BP, 315 A, 345 C, 357 G, 291 T, 0 U, 0 Other;

Query Match 100.0%; Score 1308; DB 15; Length 1308;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAGCATCTTAATCCATTTTATCTCTTGATTCGGTTAACCCCGCAATCTGCA 60
 DB 1 ATGAAAGCATCTTAATCCATTTTATCTCTTGATTCGGTTAACCCCGCAATCTGCA 60
 QY 61 TTGCTCAGAGTGAAGCCGAGCTGAAGCTGAAGTGGTGAATGTCATGTCATGAT 120
 DB 61 TTGCTCAGAGTGAAGCCGAGCTGAAGCTGAAGTGGTGAATGTCATGTCATGAT 120
 QY 121 GTGCGTGTCTCAACCAAGGCAAGCAATGATGACAGATGTCACCCAGACGATAGCCA 180
 DB 121 GTGCGTGTCTCAACCAAGGCAAGCAATGATGACAGATGTCACCCAGACGATAGCCA 180
 QY 121 GTGCGTGTCTCAACCAAGGCAAGCAATGATGACAGATGTCACCCAGACGATAGCCA 180
 DB 121 GTGCGTGTCTCAACCAAGGCAAGCAATGATGACAGATGTCACCCAGACGATAGCCA 180
 QY 181 ACCTGGCCGGTAAACTGGGTGAGCTGACACCGCGGTGGTGAAGTAACTGCTATCTC 240
 DB 181 ACCTGGCCGGTAAACTGGGTGAGCTGACACCGCGGTGGTGAAGTAACTGCTATCTC 240
 QY 181 ACCTGGCCGGTAAACTGGGTGAGCTGACACCGCGGTGGTGAAGTAACTGCTATCTC 240
 DB 181 ACCTGGCCGGTAAACTGGGTGAGCTGACACCGCGGTGGTGAAGTAACTGCTATCTC 240
 QY 241 GGACATTAATGCGGCTGAGGCTGAGTGAAGCGGATTTGCTGCTAAATGTCGCTGCCG 300
 DB 241 GGACATTAATGCGGCTGAGGCTGAGTGAAGCGGATTTGCTGCTAAATGTCGCTGCCG 300
 QY 301 CAGTCTGTGTCAGGTGCGCATTTATGCTGATGTCAGAGCGTAAACCGGTAACAGGCCGA 360
 DB 301 CAGTCTGTGTCAGGTGCGCATTTATGCTGATGTCAGAGCGTAAACCGGTAACAGGCCGA 360
 QY 361 GCTTTCGCGCGCGGCTGACCTTGAATGCTGATGTCAGAGCGTAAACCGGTAACAGGCCGA 420
 DB 361 GCTTTCGCGCGCGGCTGACCTTGAATGCTGATGTCAGAGCGTAAACCGGTAACAGGCCGA 420
 QY 421 TCCAGTCCCGATCCGTTATTTAAATCTCTAATAAACTGGGTGTCACATGATTAACCGG 480
 DB 421 TCCAGTCCCGATCCGTTATTTAAATCTCTAATAAACTGGGTGTCACATGATTAACCGG 480
 QY 481 AACGTGATGTCAGCGATCTCGAGAGGCGAGAGGCTCAATTGTCATTTACCGGGCAT 540
 DB 481 AACGTGATGTCAGCGATCTCGAGAGGCGAGAGGCTCAATTGTCATTTACCGGGCAT 540
 QY 541 TATCAAAACGCGCTTTCGCGAACTGGAACGGGTGCTTAATTTTCGCGCAATCAAACTTGTGC 600
 DB 541 TATCAAAACGCGCTTTCGCGAACTGGAACGGGTGCTTAATTTTCGCGCAATCAAACTTGTGC 600
 QY 601 CTTAAACGTGAAGAAACAGAGCAAGCTGTTAATGACGACGATTAACATCGGAATC 660
 DB 601 CTTAAACGTGAAGAAACAGAGCAAGCTGTTAATGACGACGATTAACATCGGAATC 660
 QY 661 AAGGTGAGCGCGCATGTGTCTCATTTAACCGGTGCGGTGAAGCTGCAATGCTGACG 720
 DB 661 AAGGTGAGCGCGCATGTGTCTCATTTAACCGGTGCGGTGAAGCTGCAATGCTGACG 720

QY 721 GAGATATTTCTCTGCAACAGCAACAGGAAATGCCGAGCCCGGGTGGGAAAGATCACC 780
 DB 721 GAGATATTTCTCTGCAACAGCAACAGGAAATGCCGAGCCCGGGTGGGAAAGATCACC 780
 QY 781 GATTACACACAGTGAAGACCTTGTGTAATGTTGATACAGCGCAATTTGATTTGCTACAA 840
 DB 781 GATTACACACAGTGAAGACCTTGTGTAATGTTGATACAGCGCAATTTGATTTGCTACAA 840
 QY 841 GCGACGCCAGAGTTCGCCGAGCCGCGCACCCCGTTATTAATTTGATTCACAGCAGCG 900
 DB 841 GCGACGCCAGAGTTCGCCGAGCCGCGCACCCCGTTATTAATTTGATTCACAGCAGCG 900
 QY 901 TTGAGCGCCCATCCACCCGCAAAAACAGCGCTATGTTGATTCACCTTCACTGCTG 960
 DB 901 TTGAGCGCCCATCCACCCGCAAAAACAGCGCTATGTTGATTCACCTTCACTGCTG 960
 QY 961 TTTATCGCGGACAGATTAATGTCGCAAAATCGCGGCGGCACTGAGGCTCAACTG 1020
 DB 961 TTTATCGCGGACAGATTAATGTCGCAAAATCGCGGCGGCACTGAGGCTCAACTG 1020
 QY 1021 ACGCTTCCCGGTCAAGCCGATTAACAGCGCCGCGAGTGTGAATGTTGTAACGCTG 1080
 DB 1021 ACGCTTCCCGGTCAAGCCGATTAACAGCGCCGCGAGTGTGAATGTTGTAACGCTG 1080
 QY 1081 CGTGGCTTAAGCGATTAACAGCGGATTCAGTTTCGCTGCTTCCAGCTTTACG 1140
 DB 1081 CGTGGCTTAAGCGATTAACAGCGGATTCAGTTTCGCTGCTTCCAGCTTTACG 1140
 QY 1141 CAGATGCGGTAAACCCCGCTGTCATTAATAGCCCGCGGAGGTGAACCTGAC 1200
 DB 1141 CAGATGCGGTAAACCCCGCTGTCATTAATAGCCCGCGGAGGTGAACCTGAC 1200
 QY 1201 CTGCGAGATGTAAGAGCGAAATGCGAGGCGATGTTGTTGAGGTTTACGCA 1260
 DB 1201 CTGCGAGATGTAAGAGCGAAATGCGAGGCGATGTTGTTGAGGTTTACGCA 1260
 QY 1261 ATCGTAATGAAGCAGCATACCGGCGTGCAGTTTGAATCTCATCTA 1308
 DB 1261 ATCGTAATGAAGCAGCATACCGGCGTGCAGTTTGAATCTCATCTA 1308

RESULT 4
 AED50817
 ID AED50817 standard; DNA, 1296 BP.

XX AED50817;
 AC 29-DEC-2005 (first entry)
 XX
 XX
 DE *Escherichia coli* phytase *apra* mutant DNA SEQ ID NO: 1 #1.
 XX
 KW Transgenic animal; gene therapy; feedstuff; food; thermostable; mutant;
 XX ds; coding sequence; 3-phytase; recombinant DNA; *apra* gene.
 OS *Escherichia coli*; strain K12.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1296
 FT /tag= a
 FT /EC number= "3.1.3.8"
 FT /product= "*Escherichia coli* phytase mutant protein"
 FT /note= "No stop codon"

US2005246780-A1.
 03-NOV-2005.
 11-FEB-2005; 2005US-00056354.
 13-AUG-1997; 97US-00910798.

PR 01-MAR-1999; 99US-00259214.
 PR 13-APR-1999; 99US-00291931.
 PR 25-MAY-1999; 99US-00318528.
 PR 25-MAY-2000; 2000US-00580515.
 PR 24-MAY-2001; 2001US-00863719.
 PR 24-MAY-2002; 2002US-00156660.
 XX
 PA (SHORT) SHORT J M.
 PA (KRETT) KRETT K.
 PA (GRAY) GRAY K A.
 PA (BART) BARTON N R.
 PA (GAR) GARRETT J B.
 PA (ODON) O'DONOGHUE E.
 PA (MATH) MATHUR E J.
 XX
 PI Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, Odonoghue E;
 PI Mathur EJ;
 DR WPI; 2005-733851/75.
 DR P-PSDB; AED50818.
 XX
 PT New isolated or recombinant nucleic acid encoding phytase enzymes, useful
 PT as foodstuff, and for oil degumming, producing an animal feed, delivering
 PT a phytase enzyme supplement to an animal.
 XX
 PS Claim 1; SEQ ID NO 1; 104bp; English.
 XX
 CC The present invention relates to a recombinant phytase (EC 3.1.2.8)
 CC protein and its thermostable nucleic acid molecule. Phytases of the
 CC invention is a thermostable protein. They are useful in foodstuffs, for
 CC oil degumming, producing an animal feed, delivering a phytase enzyme
 CC supplement to an animal, increasing the resistance of a phytase
 CC polypeptide to enzymatic inactivation in a digestive system of an animal
 CC and in processing of corn and sorghum kernels. The invention is also
 CC useful in gene therapy and in production of transgenic animals. The
 CC present sequence is a Escherichia coli phytase appa mutant DNA. Note: The
 CC listing. This sequence differs from the SEQ ID NO: 1 shown on page 19 in
 CC disclosure of the specification (see AED50824).
 CC
 XX
 SQ Sequence 1296 BP; 311 A; 342 C; 356 G; 287 T; 0 U; 0 Other;
 Query Match 99.1%; Score 1296; DB 14; Length 1296;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGAAAGCATCTTAATCCATTCTTCTGATTCGGTTAACCCCGCAATCTGCA 60
 DB 1 ATGAAAGCATCTTAATCCATTCTTCTGATTCGGTTAACCCCGCAATCTGCA 60
 QY 61 TTGCTCAGAGTGAAGCGGAGCTGAAGTGAAGTGTGATTCAGTCCGTCATGCT 120
 DB 61 TTGCTCAGAGTGAAGCGGAGCTGAAGTGAAGTGTGATTCAGTCCGTCATGCT 120
 QY 121 GTGCGTGTCCAAACCAAGCCAGCACTGATGACGATTCACCCAGACGATGCGCA 180
 DB 121 GTGCGTGTCCAAACCAAGCCAGCACTGATGACGATTCACCCAGACGATGCGCA 180
 QY 181 ACTGGCGGTAAACCTGGTGAAGTGAACACCGCGGCTGTGAGCTAATGCGCTATCTC 240
 DB 181 ACTGGCGGTAAACCTGGTGAAGTGAACACCGCGGCTGTGAGCTAATGCGCTATCTC 240
 QY 241 GGACATTAATGCGGTGAAGCTGTGATCCAGATTCGCTGCTAAATGTGCGTCCCG 300
 DB 241 GGACATTAATGCGGTGAAGCTGTGATCCAGATTCGCTGCTAAATGTGCGTCCCG 300
 QY 301 CAGTCTGTGACAGTCCGATTAATGCTGATTCGACGAGCTACCCGTAACAGCGCAA 360
 DB 301 CAGTCTGTGACAGTCCGATTAATGCTGATTCGACGAGCTACCCGTAACAGCGCAA 360
 QY 361 GCCTTCGCGCGCGGCTGCACTGATCTGTGCAATTAACCGTACATACCCAGCGAGATACG 420
 DB 361 GCCTTCGCGCGCGGCTGCACTGATCTGTGCAATTAACCGTACATACCCAGCGAGATACG 420

QY 421 TCCAGTCCGATCGGTTATTTAATCTCTAATAAACTGGCGTTTGCCAACTGATTAACGCG 480
 DB 421 TCCAGTCCGATCGGTTATTTAATCTCTAATAAACTGGCGTTTGCCAACTGATTAACGCG 480
 QY 481 AACGTGACTGACGCGATCTCTGAGAGGGCAGAGGGTCAATTTGCTGACCTTACCGGGCAT 540
 DB 481 AACGTGACTGACGCGATCTCTGAGAGGGCAGAGGGTCAATTTGCTGACCTTACCGGGCAT 540
 QY 541 TATCAACGCGGTTTCCGGAACCTGGAACGGGTGCTTAATTTTCGGAATCAACTGTGC 600
 DB 541 TATCAACGCGGTTTCCGGAACCTGGAACGGGTGCTTAATTTTCGGAATCAACTGTGC 600
 QY 601 CTTAAACGTGAGAAACAGACGAAAGCTGTTCAATTAACGACGCAATTACATCGAACTC 660
 DB 601 CTTAAACGTGAGAAACAGACGAAAGCTGTTCAATTAACGACGCAATTACATCGAACTC 660
 QY 661 AAGGTAGCGCCGACCTGTGTCTCATTTAACCGGTGCGTTAAGCTTCGATCAATGCTGACG 720
 DB 661 AAGGTAGCGCCGACCTGTGTCTCATTTAACCGGTGCGTTAAGCTTCGATCAATGCTGACG 720
 QY 721 GAGATATTTCTCTGCAACAGACAGAGGAATGCGGAGCCGGGGTGGGAAGATACACC 780
 DB 721 GAGATATTTCTCTGCAACAGACAGAGGAATGCGGAGCCGGGGTGGGAAGATACACC 780
 QY 781 GATTCAACCAAGTGAACACCTTGCTTAAGTTTGATTAACCGCAATTGATTTGCTACAA 840
 DB 781 GATTCAACCAAGTGAACACCTTGCTTAAGTTTGATTAACCGCAATTGATTTGCTACAA 840
 QY 841 CGCAGCGCAGAGGTTGCGCGAGCGCGGCAACCCGTTATTAAGATTGATCAAGACGCG 900
 DB 841 CGCAGCGCAGAGGTTGCGCGAGCGCGGCAACCCGTTATTAAGATTGATCAAGACGCG 900
 QY 901 TTGACGCCCCATCAACGTAATCTGGAATCTCGGCGGCGCACTGAGCTCAACTGCG 960
 DB 901 TTGACGCCCCATCAACGTAATCTGGAATCTCGGCGGCGCACTGAGCTCAACTGCG 960
 QY 961 TTTATGCGCGGACAGATTAATCTGGAATCTCGGCGGCGCACTGAGCTCAACTGCG 1020
 DB 961 TTTATGCGCGGACAGATTAATCTGGAATCTCGGCGGCGCACTGAGCTCAACTGCG 1020
 QY 1021 ACGGTTCCGCGTCAACCGGATTAACACGCGCGCAGGTGTGAACCTGTGTTGAACGCTGCG 1080
 DB 1021 ACGGTTCCGCGTCAACCGGATTAACACGCGCGCAGGTGTGAACCTGTGTTGAACGCTGCG 1080
 QY 1081 CGTGGCTAAGCGATTAACAGCCAGTGAATTCAGATTTGCTGTGCTTCCAGACTTTACAG 1140
 DB 1081 CGTGGCTAAGCGATTAACAGCCAGTGAATTCAGATTTGCTGTGCTTCCAGACTTTACAG 1140
 QY 1141 CAGATGCTGATTAACAGCCGCTGTCTTAATTAATGCGCGCGGAGAGTGAACCTGACC 1200
 DB 1141 CAGATGCTGATTAACAGCCGCTGTCTTAATTAATGCGCGCGGAGAGTGAACCTGACC 1200
 QY 1201 CTGGCAGATGTGAAGACGAAATGCGCAGGGAATGTGTTGTTGCGAGGTTTTCGCAA 1260
 DB 1201 CTGGCAGATGTGAAGACGAAATGCGCAGGGAATGTGTTGTTGCGAGGTTTTCGCAA 1260
 QY 1261 ATGCTGAATGAAGCAGCATTAACGGGCGTGCAGTTTG 1296
 DB 1261 ATGCTGAATGAAGCAGCATTAACGGGCGTGCAGTTTG 1296
 RESULT 5
 ABE75422
 ID ABE75422 standard; DNA; 1901 BP.
 AC ABE75422;
 XX
 DT 23-FEB-2006 (first entry)
 XX
 DE E. coli B modified appa phytase 819PH59 DNA sequence SEQ ID NO: 9.
 XX
 KW ds; gene; mutant; thermostable; phytase; protein stabilization;

KW pharmaceutical; fertilizer; cramp; cell culture; osteoporosis;
XX osteopathic; food; appa.
OS Escherichia coli K12.
XX Synthetic.
XX Key
FT CDS
FT Location/Qualifiers
FT 188..1486
FT /tag= a
FT /product= "modified Appa phytase"
FT replace(389,T)
FT /tag= b
FT /note= "Wild type base T replaced by G"
FT replace(390,G)
FT /tag= c
FT /note= "Wild type base G replaced by A"
FT 401..403
FT /tag= d
FT /note= "ecodes Arg"
FT replace(437,C)
FT /tag= e
FT /note= "Wild type base C replaced by T"
FT replace(438,A)
FT /tag= f
FT /note= "Wild type base A replaced by G"
FT replace(439,A)
FT /tag= g
FT /note= "Wild type base A replaced by G"
FT replace(470,G)
FT /tag= h
FT /note= "Wild type base G replaced by C"
FT replace(472,G)
FT /tag= i
FT /note= "Wild type base G replaced by T"
FT replace(476,A)
FT /tag= j
FT /note= "Wild type base A replaced by T"
FT replace(477,A)
FT /tag= k
FT /note= "Wild type base A replaced by G"
FT replace(478,G)
FT /tag= l
FT /note= "Wild type base G replaced by T"
FT replace(689,A)
FT /tag= m
FT /note= "Wild type base A replaced by G"
FT replace(690,G)
FT /tag= n
FT /note= "Wild type base G replaced by A"
FT replace(691,C)
FT /tag= o
FT /note= "Wild type base C replaced by G"
FT replace(728,C)
FT /tag= p
FT /note= "Wild type base C replaced by T"
FT replace(729,G)
FT /tag= q
FT /note= "Wild type base G replaced by A"
FT replace(730,G)
FT /tag= r
FT /note= "Wild type base G replaced by T"
FT replace(863,A)
FT /tag= s
FT /note= "Wild type base A replaced by T"
FT replace(864,A)
FT /tag= t
FT /note= "Wild type base A replaced by G"
FT replace(1016,T)
FT /tag= u
FT /note= "Wild type base T replaced by G"
XX US2005281792-A1.

PD 22-DEC-2005.
XX
XX 01-SEP-2004; 2004US-00933115.
XX
XX 13-AUG-1997; 97US-00910798.
PR 13-MAR-1999; 99US-00259214.
PR 13-APR-1999; 99US-00291931.
PR 25-MAY-1999; 99US-00318528.
PR 25-MAY-2000; 2000US-00580515.
PR 24-MAY-2001; 2001US-00866379.
XX
XX (SHOR/) SHORT J M.
PA (KRETT/) KRETTZ K A.
PA (GRAY/) GRAY K A.
PA (BART/) BARTON N R.
PA (GARR/) GARRETT J B.
PA (ODON/) ODOENGHUE E.
PA (BAUM/) BAUM W.
PA (ROBE/) ROBERTSON D B.
PA (ZORN/) ZORNER P.
XX
XX Short JM, Krettz KA, Gray KA, Barton NR, Garrett JB, Odonoghue E;
PI Baum W, Robertson DE, Zorner P;
XX
XX MPI: 2006-055022/06.
XX P-PSDB; ABE75421.
XX
XX Formulation useful as dietary supplement for treating, preventing or
PT reversing osteoporosis and bone loss, and preventing muscle cramps,
PT comprises one or more polypeptides having phytase activity.
XX
XX Claim 1; Page: 82pp; English.
XX
XX The present sequence is that of a polynucleotide encoding a novel
CC modified phytase derived from the Escherichia coli appa protein, with
CC improved thermal tolerance and protease stability compared to the wild-
CC type. The present invention relates to a novel formulation, comprising at
CC least one polypeptide having phytase activity, where the polypeptide is a
CC modified sequence (ABE75421) derived from the Escherichia coli K-12 appa
CC phytase (ABE75419). The modification of the enzyme, comprising a series
CC of mutations, improves the thermal tolerance and protease stability of
CC the protein. The specification also claims a pharmaceutical composition;
CC a kit; an immobilized phytase; a fertilizer or soil additive; a liquid
CC supplement for preventing muscle cramps; a hydrating agent; a tissue
CC culture or cell culture media; and a plant food additive, all comprising
CC the novel phytase, and a method of reducing pollution and increasing
CC nutrient availability in an environment or environmental sample by
CC degrading environmental phytic acid. The novel enzyme is a dietary
CC supplement useful for treating, preventing or reversing osteoporosis or
CC bone loss, and preventing muscle cramps. The liquid supplement is useful
CC for preventing muscle cramps. The method is useful for reducing pollution
CC and increasing nutrient availability in an environment or environmental
CC sample by degrading environmental phytic acid, where the environment or
CC environmental sample comprises a soil or a body of water. The immobilized
CC phytase is useful in foodstuffs for improving the feeding value of
CC phyate rich ingredients. NOTE: This sequence is not shown in the
CC specification but was created from the wild-type sequence (ABE75419) and
XX the information given in claim 1.
XX
XX Sequence 1901 BP; 470 A; 497 C; 500 G; 433 T; 0 U; 1 Other;
SQ
Query Match 98.9%; Score 1294.2; DB 15; Length 1901;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGAAGGAGCTTAATCCATTCTTCTGATTCGGTAACCCGCAATCGCA 60
DB 188 ATGAAGGAGCTTAATCCATTCTTCTGATTCGGTAACCCGCAATCGCA 247
QY 61 TTCGCTCAGAGTGAGCCGAGAGCTGAGAAAGTGATGATTCAGTCGATGAT 120
DB 248 TTCGCTCAGAGTGAGCCGAGAGCTGAGAAAGTGATGATTCAGTCGATGAT 307

QY 121 GTGGTGCTCCAAAGCAAGCCCAACTGATGAGAGATGTCAACCCAGACGATGGCA 180
Db 308 GTGGTGCTCCAAAGCAAGCCCAACTGATGAGAGATGTCAACCCAGACGATGGCA 367
QY 181 ACCTGGCCGGTAAACTGGGTAGCTGACACCGCGGTGGTGAAGCTAATGCTATCTC 240
Db 368 ACCTGGCCGGTAAACTGGGTAGCTGACACCGCGGTGGTGAAGCTAATGCTATCTC 427
QY 241 GGAGATTACTGGCGCTGACCGCTGTGTAGCCGACGAGATTGCTCTAAATGTGGTCCCG 300
Db 428 GGAGATTACTGGCGCGCCGCTGTGTAGCCGACGAGATTGCTCTAAATGTGGTCCCG 487
QY 301 CAGTCTGGTCAAGTTCGCGATTATGTGATGTGACAGAGCTACCCGTTAAACAGCGAA 360
Db 488 CAGTCTGGTCAAGTTCGCGATTATGTGATGTGACAGAGCTACCCGTTAAACAGCGAA 547
QY 361 GCCTTCGCGCGCGGCTGGACCTGACTGTGCAATTAACCGTACATACCAGAGATACG 420
Db 548 GCCTTCGCGCGCGGCTGGACCTGACTGTGCAATTAACCGTACATACCAGAGATACG 607
QY 421 TCCAGTCCCGATCCGTTATTTAAATCCTCTAAAACTGGCGTTGGCCAACTGGATACGG 480
Db 608 TCCAGTCCCGATCCGTTATTTAAATCCTCTAAAACTGGCGTTGGCCAACTGGATACGG 667
QY 481 AACGTGACTGACGGATCTCTGAGAGGCGACAGAGGCTCAATTGCTGACTTTACCGGCA 540
Db 668 AACGTGACTGACGGATCTCTGAGAGGCGACAGAGGCTCAATTGCTGACTTTACCGGCA 727
QY 541 TATCAACGCGGTTTCGGAACCTGGAACGGGTGCTTAAATTTTCGCAATCAACTGTGC 600
Db 728 TATCAACGCGGTTTCGGAACCTGGAACGGGTGCTTAAATTTTCGCAATCAACTGTGC 787
QY 601 CTTTAAACGTGAGAAACAGGAAGAAAGCTTCAATTAACGAGGCAATTAACGGAATC 660
Db 788 CTTTAAACGTGAGAAACAGGAAGAAAGCTTCAATTAACGAGGCAATTAACGGAATC 847
QY 661 AAGGTGAGCGCCGACTGTGTCTCATTTAACCGGTGCGTAACTCTGATCATGCTGACG 720
Db 848 AAGGTGAGCGCCGACTGTGTCTCATTTAACCGGTGCGTAACTCTGATCATGCTGACG 907
QY 721 GAGATATTTCTCTGCAACAGACACAGGAATGCGGAGCGGGGTGGGAAAGATCAC 780
Db 908 GAGATATTTCTCTGCAACAGACACAGGAATGCGGAGCGGGGTGGGAAAGATCAC 967
QY 781 GATTCAACAGGAGTGAACCTTGTGAATTTGATTAACGCGCAATTTGATTTCTACAA 840
Db 968 GATTCAACAGGAGTGAACCTTGTGAATTTGATTAACGCGCAATTTGATTTCTACAA 1027
QY 841 CGCAGCGCAGAGTGGCCCGCAGCGCGCACCCGTTATTAAGATTGATCAAGACGCG 900
Db 1028 CGCAGCGCAGAGTGGCCCGCAGCGCGCACCCGTTATTAAGATTGATCAAGACGCG 1087
QY 901 TTGACGCCCCATCAACCGCAAAAACAGGCGTATGATGATTAACCACTTCAGTGTG 960
Db 1088 TTGACGCCCCATCAACCGCAAAAACAGGCGTATGATGATTAACCACTTCAGTGTG 1147
QY 961 TTTATTCGCGGACACAGATTAATCTTGCAAAATCCGCGCGCGACCTGAGCTCAACTG 1020
Db 1148 TTTATTCGCGGACACAGATTAATCTTGCAAAATCCGCGCGCGACCTGAGCTCAACTG 1207
QY 1021 ACGCTTCCCGGTCAACCGGATACACGCGCCAGGTGTGAATGTGTGGAACGCTGCG 1080
Db 1208 ACGCTTCCCGGTCAACCGGATACACGCGCCAGGTGTGAATGTGTGGAACGCTGCG 1267
QY 1081 CGTGGGCTAAGCGATTAACGCGCACTGATTCAGGTTTCGCTGCTTCAGACTTTACG 1140
Db 1268 CGTGGGCTAAGCGATTAACGCGCACTGATTCAGGTTTCGCTGCTTCAGACTTTACG 1327
QY 1141 CAGATGCGTGAATAAACCGCTGTCTATTAATACGCGCCCGAGAGGTGAATACTGACC 1200
Db 1328 CAGATGCGTGAATAAACCGCTGTCTATTAATACGCGCCCGAGAGGTGAATACTGACC 1387
QY 1201 CTGACAGATGTGAAGAGCAAAATCGCAGAGGCGATGTGTGCTGTCAGGTTTACGCA 1260

Db 1388 CTGGCAGGATGTGAAGAGCAAAATGCGCAGGATGTGTGTCAGGTTTACGCA 1447
QY 1261 ATCGTGAATGAACGACGATACCGGCGTGCAGTTTGAGAT 1300
Db 1448 ATCGTGAATGAACGACGATACCGGCGTGCAGTTTGAT 1487
RESULT 6
AD050296 standard; DNA; 1901 BP.
XX
AC AD050296;
XX
DF 29-JUN-2004 (first entry)
XX
DE Escherichia coli K12 apba phytase mutant DNA #2.
XX
KW Apba phytase; bacteria; thermal tolerance; protease stability; foodstuff;
KM phytate; animal feed; fish feed; dough; baking; gene; mutant; ds.
XX
OS Escherichia coli; K12.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT mutation /*tag= a
FT mutation /*tag= b
FT mutation /*tag= c
XX
PN US2004091968-A1.
XX
PD 13-MAY-2004.
XX
PF 20-JUN-2003; 2003US-00601319.
XX
PR 13-AUG-1997; 97US-00910798.
PR 01-MAR-1999; 99US-00259214.
PR 13-APR-1999; 99US-00291931.
PR 25-MAY-1999; 99US-00318528.
PR 25-MAY-2000; 2000US-00580515.
PR 24-MAY-2001; 2001US-00866379.
XX
PA (SHOR/) SHORT J M.
PA (KRET/) KRETZ K.
PA (GRAY/) GRAY K A.
PA (BART/) BARTON N R.
PA (GAR/) GARRETT J B.
PA (ODON/) O'DONOGHUE E.
PA (MATH/) MATHER E J.
XX
PI Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI Mather Ed;
XX
DR WPI; 2004-374952/35.
XX
PT Producing phytase, involves providing nucleic acid derived from bacteria
PT encoding polypeptide having phytase activity, and expressing nucleic acid
PT in yeast.
XX
PS Disclosure; SEQ ID NO 6; 74pp; English.
XX
XX The invention relates to a method for producing phytase that involves
CC providing a nucleic acid encoding phytase derived from a bacteria, and
CC expressing the nucleic acid in a yeast under conditions that allow
CC expression of the enzyme in the yeast. The invention also relates to
CC modified phytase enzyme which has improved thermal tolerance and protease
CC stability at low pH. The phytase enzyme can be used in foodstuffs to
CC improve the feeding value of phytate rich ingredients, and in diet of
CC numerous animals including mammals, fowls and fishes, commercially
CC significant mammals such as pigs, goats, laboratory rodents, commercially

CC significant avian species such as chicken, ducks, doves, parrot, etc.,
 CC commercially farmed fish such as guppy, zebrafish, molly, swordtail.
 CC etc., in dough making and baking, in dietary aids for animals. The method
 CC provides easy manufacture of the active ingredient loaded biocompatible
 CC composition, higher yields and loading efficiency. The phytase
 CC incorporated in to the dietary aid is safe for animals. The present
 CC sequence is *Escherichia coli* K12 *apra* phytase mutant DNA.

XX
 SQ Sequence 1901 BP; 472 A; 498 C; 501 G; 429 T; 0 U; 1 Other;

Query Match 97.0%; Score 1268.6; DB 12; Length 1901;
 Best Local Similarity 98.5%; Pred. No. 0;
 Matches 1280; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 ATGAAAGGATCTTATCCCATTTTATCTCTTGTATCCGTAAACCCCGCATTTGCA 60
 DB 188 ATGAAAGGATCTTATCCCATTTTATCTCTTGTATCCGTAAACCCCGCATTTGCA 247
 QY 61 TTGCTCAGAGTGAAGCCGAGCTGAAGTGAAGTGTGTATGTCTAGTCTCATGTGT 120
 DB 248 TTGCTCAGAGTGAAGCCGAGCTGAAGTGAAGTGTGTATGTCTAGTCTCATGTGT 307
 QY 121 GTGCTGTCTCAACCAAGGCAAGCACTGATGACAGATGTCAACCCAGACGCATGGCA 180
 DB 308 GTGCTGTCTCAACCAAGGCAAGCACTGATGACAGATGTCAACCCAGACGCATGGCA 367
 QY 181 ACCTGGCCGCTTAACTGGGTGAGTGAACACCGCGGGTGTGTATGTCTATGTCTC 240
 DB 368 ACCTGGCCGCTTAACTGGGTGAGTGAACACCGCGGGTGTGTATGTCTATGTCTC 427
 QY 241 GGACATTAATGAGCTGACAGGCTGTGTGAGCGACGAGTTGCTGCTAAATGTGTGCGCG 300
 DB 428 GGACATTAATGAGCTGACAGGCTGTGTGAGCGACGAGTTGCTGCTAAATGTGTGCGCG 487
 QY 301 CAGTCTGTGAGTGTGCGATTATGTGTGATGTGACAGACGTAACCCGTAAACAGGCGAA 360
 DB 488 CAGTCTGTGAGTGTGCGATTATGTGTGATGTGACAGACGTAACCCGTAAACAGGCGAA 547
 QY 361 GCTTTCGCGCGCGGCTGCGACCTGATCTGTGCAATACCTGACATACCCAGGCAATAG 420
 DB 548 GCTTTCGCGCGCGGCTGCGACCTGATCTGTGCAATACCTGACATACCCAGGCAATAG 607
 QY 421 TCCAGTCCCGATCCGTTATTTATCTCTTAAACCTGGGCTTGGCAATGTGATACGCG 480
 DB 608 TCCAGTCCCGATCCGTTATTTATCTCTTAAACCTGGGCTTGGCAATGTGATACGCG 667
 QY 481 AACGTGACTGACGCGATCTCGAGAGGCGAGGAGGCTCAATTGCTGACTTTACCGGCGAT 540
 DB 668 AACGTGACTGACGCGATCTCGAGAGGCGAGGAGGCTCAATTGCTGACTTTACCGGCGAT 727
 QY 541 TATCAAAGCGGCTTTCGGAACCTGGAACGGGTCTTAAATTTTCGCAATCAAACTTTGTC 600
 DB 728 CGGCAAAAGCGGCTTTCGGAACCTGGAACGGGTCTTAAATTTTCGCAATCAAACTTTGTC 787
 QY 601 CTTAAACGTGAGAAACAGAGAGAAAGCTGTTCAATTAAGCAGGCACTTACATCGGAATC 660
 DB 788 CTTAAACGTGAGAAACAGAGAGAAAGCTGTTCAATTAAGCAGGCACTTACATCGGAATC 847
 QY 661 AAGGTAGCGCGCATGTGTCTCATTTAACCGGTGCGGTAAAGCTGCAATCATGTCTGACG 720
 DB 848 AAGGTAGCGCGCATGTGTCTCATTTAACCGGTGCGGTAAAGCTGCAATCATGTCTGACG 907
 QY 721 GAGATATTTCTCTGCAACAGACAGGGAATGCGGAGCCGGGGTGGGAAAGATCAC 780
 DB 908 GAGATATTTCTCTGCAACAGACAGGGAATGCGGAGCCGGGGTGGGAAAGATCAC 967
 QY 781 GATTCAACAGTGAACCACTTGTCAAGTTTGAACGCGCAATTTGATTGTCTACAA 840
 DB 968 GATTCAACAGTGAACCACTTGTCTAAGTTTGAACGCGCAATTTTATTGTCTACAA 1027
 QY 841 CGCAGCGCAAGAGTTGCGCGACGCGCGCAACCCGTTATTTAGATTGTGATCAAGACAGCG 900
 DB 1028 CGCAGCGCAAGAGTTGCGCGACGCGCGCAACCCGTTATTTAGATTGTGATCAAGACAGCG 1087

QY 901 TTGAGCCCCCATTCACCGGCAAAAACAGGCGTATGCTGTGACATTAACCACTTCAAGTCTG 960
 DB 1088 TTGAGCCCCCATTCACCGGCAAAAACAGGCGTATGCTGTGACATTAACCACTTCAAGTCTG 1147
 QY 961 TTTATCGCGGACACAGATTAATCTGTGCAAAATCTCGCGCGGACATGAGCTCAACTGG 1020
 DB 1148 TTTATCGCGGACACAGATTAATCTGTGCAAAATCTCGCGCGGACATGAGCTCAACTGG 1207
 QY 1021 ACGTTCCCGGTACCGCGGATTAACGCGCGCAGGTGTGAATCTGTGTTGAACGCTGG 1080
 DB 1208 ACGTTCCCGGTACCGCGGATTAACGCGCGCAGGTGTGAATCTGTGTTGAACGCTGG 1267
 QY 1081 CGTGGCTTAAAGGATTAACAGCCAGTGAATTCAGTTTGTGCTGTCTTCCAGCTTTACAG 1140
 DB 1268 CGTGGCTTAAAGGATTAACAGCCAGTGAATTCAGTTTGTGCTGTCTTCCAGCTTTACAG 1327
 QY 1141 CAGATGCGTGAATTAACCCGCTGTCAATTAATACGCGCGCGGAGAGGTGAACCTGACC 1200
 DB 1328 CAGATGCGTGAATTAACCCGCTGTCAATTAATACGCGCGCGGAGAGGTGAACCTGACC 1387
 QY 1201 CTGGCAGATGTGAAGACCGAAATGCGCAGGCGATGTGTTCTGTGGCAGAGTTTACGCA 1260
 DB 1388 CTGGCAGATGTGAAGACCGAAATGCGCAGGCGATGTGTTCTGTGGCAGAGTTTACGCA 1447
 QY 1261 ATCGTGAATGAAGCAGCATACCGGCGTGCATTTGAAT 1300
 DB 1448 ATCGTGAATGAAGCAGCATACCGGCGTGCATTTGAAT 1487

RESULT 7

AD050295
 ID AD050295 standard; DNA, 1901 BP.

XX
 AC AD050295;

XX
 DT 29-JUL-2004 (first entry)

XX
 DE *Escherichia coli* K12 *apra* phytase mutant DNA #1.

XX
 KW *Apna* phytase; bacteria; thermal tolerance; protease stability; foodstuff;

XX
 KM phytase; animal feed; fish feed; dough; baking; gene; mutant; de.

XX
 OS *Escherichia coli*; K12.

XX
 OS Synthetic.

XX
 FH Key Location/Qualifiers

XX
 FT mutation replace(389,T)

XX
 FT mutation /*tag= a

XX
 FT mutation /*tag= b

XX
 PD US2004091968-A1.

XX
 PF 13-MAY-2004.

XX
 PF 20-JUN-2003; 2003US-00601319.

XX
 PR 13-AUG-1997; 97US-00910798.

XX
 PR 01-MAR-1999; 99US-00259214.

XX
 PR 13-APR-1999; 99US-00291931.

XX
 PR 25-MAY-1999; 99US-00318528.

XX
 PR 25-MAY-2000; 2000US-00580515.

XX
 PR 24-MAY-2001; 2001US-00866379.

XX
 PA (SHORT) SHORT J M.

XX
 PA (KRET) KRETZ K.

XX
 PA (GRAY) GRAY K A.

XX
 PA (BART) BARTON N R.

XX
 PA (GAR) GARRETT J B.

XX
 PA (ODON) O'DONOGHUE E.

XX
 PA (MATH) MATHER B J.

PI Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI Mather EJ;
XX WPI: 2004-374952/35.
XX Producing phytase, involves providing nucleic acid derived from bacteria
PT encoding polypeptide having phytase activity, and expressing nucleic acid
PT in yeast.
XX
XX Disclosure; SEQ ID NO 5; 74pp; English.
XX
XX The invention relates to a method for producing phytase that involves
CC providing a nucleic acid encoding phytase derived from a bacteria, and
CC expressing the nucleic acid in a yeast under conditions that allow
CC expression of the enzyme in the yeast. The invention also relates to
CC modified phytase enzyme which has improved thermal tolerance and protease
CC stability at low pH. The phytase enzyme can be used in foodstuffs to
CC improve the feeding value of phytate rich ingredients, and in diet of
CC numerous animals including mammals, fowls and fishes, commercially
CC significant mammals such as pigs, goats, laboratory rodents, commercially
CC significant avian species such as chicken, ducks, doves, parrot, etc.,
CC commercially farmed fish such as supply, zebrafish, molly, swordtail,
CC etc., in dough making and baking, in dietary aids for animals. The method
CC provides easy manufacture of the active ingredient loaded biocompatible
CC composition, higher yields and loading efficiency. The phytase
CC incorporated in to the dietary aid is safe for animals. The present
CC sequence is Escherichia coli K12 apba phytase mutant DNA.
XX
XX Sequence 1901 BP; 475 A; 499 C; 499 G; 427 T; 0 U; 1 Other;
SO
Query Match 96.9%; Score 1267; DB 12; Length 1901;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 1279; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 1 ATGAAGGAGCTTATCCATTTTATCTTCTGATTCGGTTAACCCCGCATCTGCA 60
DB 188 ATGAAGGAGCTTATCCATTTTATCTTCTGATTCGGTTAACCCCGCATCTGCA 247
QY 61 TTGCTCAGATGAGCGGAGCTGAAAGTGAGTGTGATGTCAGTGCATGAT 120
DB 248 TTGCTCAGATGAGCGGAGCTGAAAGTGAGTGTGATGTCAGTGCATGAT 307
QY 121 GTGCGTGTCCAAACCAAGCCAGCAGCACTGATGTCAGATGTCAGATGTCAG 180
DB 308 GTGCGTGTCCAAACCAAGCCAGCAGCACTGATGTCAGATGTCAGATGTCAG 367
QY 181 ACTGCGCGGTAAACCTGGGTGAGTGAACACCGCGGTGTGATGATGCTATCTC 240
DB 368 ACTGCGCGGTAAACCTGGGTGAGTGAACACCGCGGTGTGATGATGCTATCTC 427
QY 241 GGAACATTACGCGGTGAGGTGCTGAGCGAGATGCTGCGTAAATGAGCGTGC 300
DB 428 GGAACATTACGCGGTGAGGTGCTGAGCGAGATGCTGCGTAAATGAGCGTGC 487
QY 301 CAGTCTGTGAGTGCAGATTTATGCTGATGTCAGACGAGTACCCGTTAAACAGGCGAA 360
DB 488 CAGTCTGTGAGTGCAGATTTATGCTGATGTCAGACGAGTACCCGTTAAACAGGCGAA 547
QY 361 GCTTTCGCGCGCGGCTGACCTGATGTCGATTAACCGTACATACCCAGCGAGATACG 420
DB 548 GCTTTCGCGCGCGGCTGACCTGATGTCGATTAACCGTACATACCCAGCGAGATACG 607
QY 421 TCCAGTCCCGATCCGTTATTTAAATCTCTAAATAATGCGGTGCACTGATTAACGCG 480
DB 608 TCCAGTCCCGATCCGTTATTTAAATCTCTAAATAATGCGGTGCACTGATTAACGCG 667
QY 481 AACGTGACGTGACGCGATCTCGAGAGGCGAGAGGCTCAATTGCTTAACCGGCGAT 540
DB 668 AACGTGACGTGACGCGATCTCGAGAGGCGAGAGGCTCAATTGCTTAACCGGCGAT 727
QY 541 TATCAACGCGGCTTTCGCGAACTGGAACGGGTGCTTAATTTTCGCGAATCAACTTGTGC 600
DB 728 CGGCAACGCGGCTTTCGCGAACTGGAACGGGTGCTTAATTTTCGCGAATCAACTTGTGC 787

QY 601 CTTAAACGTGAGAAACAGACGAAAGCTGTTCAATTAACGAGGCAATACCATCGAACTC 660
DB 788 CTTAAACGTGAGAAACAGACGAAAGCTGTTCAATTAACGAGGCAATACCATCGAACTC 847
QY 661 AAGGTAGCGCCGCACTGTGTCTCAATTAACCGGTGCGGTAAAGCTCGCATTAATGTCACG 720
DB 848 AAGGTAGCGCCGCACTGTGTCTCAATTAACCGGTGCGGTAAAGCTCGCATTAATGTCACG 907
QY 721 GAGATATTTCTCTGCAACAAAGCAAGGGAATGCGGAGCGGGGTGGGAAAGATACCC 780
DB 908 GAGATATTTCTCTGCAACAAAGCAAGGGAATGCGGAGCGGGGTGGGAAAGATACCC 967
QY 781 GATTTCACACAGTGAACACCTTGCTAATGATTAACCGGCAATTTGATTTGCTACAA 840
DB 968 GATTTCACACAGTGAACACCTTGCTAATGATTAACCGGCAATTTGATTTGCTACAA 1027
QY 841 CGCAGCGCAAGGTTGCGCGCAGCGCCGCAACCCGTTAATTAATTTGATCAAGACGCG 900
DB 1028 CGCAGCGCAAGGTTGCGCGCAGCGCCGCAACCCGTTAATTAATTTGATCAAGACGCG 1087
QY 901 TTGACGCGCATCCACCGCAAAACAGCGGTATGTTGATTAATCCCATTCAGTGC 960
DB 1088 TTGACGCGCATCCACCGCAAAACAGCGGTATGTTGATTAATCCCATTCAGTGC 1147
QY 961 TTATCCCGCGACAGATTAATCTGCAAAATCTCGCGCGCACTGAGCTCAACTCG 1020
DB 1148 TTATCCCGCGACAGATTAATCTGCAAAATCTCGCGCGCACTGAGCTCAACTCG 1207
QY 1021 ACGTTTCCCGGTGACCGCGATTAACGCGCGCAGGTGTGAATGTTGTAACGCTCG 1080
DB 1208 ACGTTTCCCGGTGACCGCGATTAACGCGCGCAGGTGTGAATGTTGTAACGCTCG 1267
QY 1081 CGTGGCTAAGCAGTAAACAGCAGTGAATCAGGTTGCGTGTGCTTCAAGCTTACAG 1140
DB 1268 CGTGGCTAAGCAGTAAACAGCAGTGAATCAGGTTGCGTGTGCTTCAAGCTTACAG 1327
QY 1141 CAGATGCGTATTAACGCGCTGTCTTAATTAACGCGCGCGAGAGTGAACCTGACC 1200
DB 1328 CAGATGCGTATTAACGCGCTGTCTTAATTAACGCGCGAGAGTGAACCTGACC 1387
QY 1201 CTGCGAGATGTGAAGAGCGAATATGCGCAGGCGATGTTGCTGTCAGGTTTACGCAA 1260
DB 1388 CTGCGAGATGTGAAGAGCGAATATGCGCAGGCGATGTTGCTGTCAGGTTTACGCAA 1447
QY 1261 ATCGTAATGAAGCAGCATACCGGCGTGCAGTTTGAT 1300
DB 1448 ATCGTAATGAAGCAGCATACCGGCGTGCAGTTTGAT 1487
RESULT 8
ABE75416
ID ABE75416 standard; DNA, 1901 BP.
XX
XX ABE75416;
AC
XX
DT 23-FEB-2006 (first entry)
XX
DE E. coli phytase related DNA sequence SEQ ID NO: 5.
XX
KW ds; thermostable; phytase; protein stabilization; pharmaceutical;
XX fertilizer; cramp; cell culture; osteoporosis; osteopathic; food.
OS Escherichia coli.
XX
PN US2005281792-A1.
PD
XX
PD 22-DEC-2005.
PF
XX 01-SEP-2004; 2004US-0093115.
XX
XX 13-AUG-1997; 97US-00910798.
PR
XX 01-MAR-1999; 99US-00259214.

PR 13-APR-1999; 99US-00291931.
 PR 25-MAY-1999; 99US-00318528.
 PR 25-MAY-2000; 2000US-00580515.
 PR 24-MAY-2001; 2001US-00866379.
 XX
 PA (SHOR/) SHORT J M.
 PA (KRETZ/) KRETZ K A.
 PA (GRAY/) GRAY K A.
 PA (BART/) BARTON N R.
 PA (GAR/) GARRETT J B.
 PA (ODON/) ODOMOSHUB E.
 PA (BAUM/) BAUM W.
 PA (ROBE/) ROBERTSON D E.
 PA (ZORN/) ZORNER P.
 XX
 PI Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, Odomoshue E;
 PI Baum W, Robertson DE, Zorner P;
 DR WPI; 2006-055022/06.
 XX
 PT Formulation useful as dietary supplement for treating, preventing or
 PT reversing osteoporosis and bone loss, and preventing muscle cramps,
 PT comprising one or more polypeptides having phytase activity.
 XX
 PS Disclosure; SEQ ID NO 5; 82pp; English.
 XX
 CC The present sequence is a DNA sequence related to the production of the
 CC novel thermolabile / protease resistant *Bacillus coli* phytase
 CC described in the specification. The present invention relates to a novel
 CC formulation, comprising at least one polypeptide having phytase activity,
 CC where the polypeptide is a modified sequence (ABE75421) derived from the
 CC *Bacillus coli* K-12 *appa* phytase (ABE75419). The modification of the
 CC enzyme, comprising a series of mutations, improves the thermal tolerance
 CC and process stability of the protein. The specification also claims a
 CC pharmaceutical composition; a kit; an immobilized phytase; a fertilizer
 CC or soil additive; a liquid supplement for preventing muscle cramps; a
 CC hydrating agent; a tissue culture or cell culture media; and a plant food
 CC additive, all comprising the novel phytase, and a method of reducing
 CC pollution, and increasing nutrient availability in an environment or
 CC environmental sample by degrading environmental phytic acid. The novel
 CC enzyme is a dietary supplement useful for treating, preventing or
 CC reversing osteoporosis or bone loss, and preventing muscle cramps. The
 CC liquid supplement is useful for preventing muscle cramps. The method is
 CC useful for reducing pollution and increasing nutrient availability in an
 CC environment or environmental sample by degrading environmental phytic
 CC acid, where the environment or environmental sample comprises a soil or a
 CC body of water. The immobilized phytase is useful in foodstuffs for
 CC improving the feeding value of phytate rich ingredients. NOTES: The
 CC present sequence is shown in the sequence listing but not discussed
 CC elsewhere in the specification.
 CC
 XX
 SQ Sequence 1901 BP; 475 A; 499 C; 499 G; 427 T; 0 U; 1 Other;
 Query Match 96.9%; Score 1267; DB 15; Length 1901;
 Best Local Similarity 98.4%; Pred. No. 0;
 Matches 1279; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
 QY 1 ATGAAGGATCTTAATCCATTTTATCTTTGATTCGTTAACCCCGCAATCTGCA 60
 DB 188 ATGAAGGATCTTAATCCATTTTATCTTTGATTCGTTAACCCCGCAATCTGCA 247
 QY 61 TTGCGTCAAGTGAAGCCGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 120
 DB 248 TTGCGTCAAGTGAAGCCGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 307
 QY 121 GTGCGTCTCAACCAAGCCGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 180
 DB 308 GTGCGTCTCAACCAAGCCGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 367
 QY 181 ACTGGCCGGTAAACTGGTGAAGTGAACACCGCGCGTGTGAAGTGAAGTGAAGT 240
 DB 368 ACTGGCCGGTAAACTGGTGAAGTGAACACCGCGCGTGTGAAGTGAAGTGAAGT 427

QY 241 GGAACATTAATGCGGTCAGGTCGTGTAGCCGCAATGTCGCTTAATGTCGTCGCCG 300
 DB 428 GGAACATTAATGCGGTCAGGTCGTGTAGCCGCAATGTCGCTTAATGTCGTCGCCG 487
 QY 301 CAGTCTGTGTCAGGTCGCAATTAATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 360
 DB 488 CAGTCTGTGTCAGGTCGCAATTAATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 547
 QY 361 GCTTCG 420
 DB 548 GCTTCG 607
 QY 421 TCCAGTCCCGATCCGCTTAATTAATCCCTTAATAAATGCGCTTGGCACTGATTAAC 480
 DB 608 TCCAGTCCCGATCCGCTTAATTAATCCCTTAATAAATGCGCTTGGCACTGATTAAC 667
 QY 481 AACGTGATGACGCGATCTCGAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
 DB 668 AACGTGATGACGCGATCTCGAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 727
 QY 541 TATCAACGCGGCTTTCGCAACTGGAACGCGGTCCTTAATTTTCGCAATCAAACTT 600
 DB 728 CGGCAAAAGCGGCTTTCGCAACTGGAACGCGGTCCTTAATTTTCGCAATCAAACTT 787
 QY 601 CTTAAACGTGAGAAACAGAAAGGCTTCAATTAACGCGGATTAACGCGGATTAAC 660
 DB 788 CTTAAACGTGAGAAACAGAAAGGCTTCAATTAACGCGGATTAACGCGGATTAAC 847
 QY 661 AAGGTGACG 720
 DB 848 AAGGTGACG 907
 QY 721 GAGATATTTCTCTGCAACAGCAAGGAAATGCGGACCGCGGTCGGAAGATCAAC 780
 DB 908 GAGATATTTCTCTGCAACAGCAAGGAAATGCGGACCGCGGTCGGAAGATCAAC 967
 QY 781 GATTCAACGCGATGGAACACCTTGTGAATTTGCAATACGCGCAATTTGATTTGCTA 840
 DB 968 GATTCAACGCGATGGAACACCTTGTGAATTTGCAATACGCGCAATTTGATTTGCTA 1027
 QY 841 CGCAGCGCAAGGTTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
 DB 1028 CGCAGCGCAAGGTTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1087
 QY 901 TTGACGCGCCATCAACCGCAAAACAGCGGTATGCTGATTAACCACTTCACTGCTG 960
 DB 1088 TTGACGCGCCATCAACCGCAAAACAGCGGTATGCTGATTAACCACTTCACTGCTG 1147
 QY 961 TTTATCGCGCGCAACGATTAATCTGCAAAATCTCGCGCGCGCGCGCGCGCGCGCG 1020
 DB 1148 TTTATCGCGCGCAACGATTAATCTGCAAAATCTCGCGCGCGCGCGCGCGCGCGCG 1207
 QY 1021 ACGCTTCCCGGTGACCGGATTAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
 DB 1208 ACGCTTCCCGGTGACCGGATTAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1267
 QY 1081 CGTGGCTTAAGCGATTAACAGCGCGATTAATCAAGTTTCGCTGCTTCCAGACTTACAG 1140
 DB 1268 CGTGGCTTAAGCGATTAACAGCGCGATTAATCAAGTTTCGCTGCTTCCAGACTTACAG 1327
 QY 1141 CAGATGCGGTATTAACGCGCGCTTCAATTAATACGCGCGCGCGCGCGCGCGCGCG 1200
 DB 1328 CAGATGCGGTATTAACGCGCGCTTCAATTAATACGCGCGCGCGCGCGCGCGCGCG 1387
 QY 1201 CTGGCAGGATGTAAGGCGGAAATGCGCAGGCGCGGATGCTGTTGAGAGGTTTAACGCA 1260
 DB 1388 CTGGCAGGATGTAAGGCGGAAATGCGCAGGCGCGGATGCTGTTGAGAGGTTTAACGCA 1447
 QY 1261 ATCGTGAATGAAGCAGCATACCGCGCGTGAAGTTTGAT 1300
 DB 1448 ATCGTGAATGAAGCAGCATACCGCGCGTGAAGTTTGAT 1487

RESULT 9
AAC8885
ID AAC8885 standard; DNA; 1323 BP.
AC AAC8885;
XX
XX 07-MAR-2001 (first entry)
XX
XX Escherichia coli B phytase enzyme nucleotide sequence.
DE
XX Escherichia coli B; phytase enzyme; anabolic; phytate digestion;
KM nutrition; db.
XX
XX Escherichia coli.
OS
XX W0200071728-A1.
XX
XX 30-NOV-2000.
XX
XX 25-MAY-2000; 2000MO-US014846.
XX
XX 25-MAY-1999; 9905-00318528.
XX
XX (DIVE-) DIVERSA CORP.
XX
XX Short JM, Kretz KA;
XX
XX WPI; 2001-112081/12.
XX
XX P-PSDB; AAB37892.
XX
XX Improving the nutritional value of phytate-containing foodstuffs, using
PT phytase enzymes which catalyze the liberation of inorganic phosphate from
PT the phytates.
XX
XX Claim 2; Fig 1; 147bp; English.
XX
XX The present sequence encodes a phytase enzyme from Escherichia coli B.
CC The enzyme catalyses the liberation of inorganic phosphate from the
CC phytate in phytate-containing foodstuffs and can thus be used to improve
CC the nutritional value of phytate rich ingredients
XX
SQ Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;
Query Match 96.7%; Score 1265; DB 5; Length 1323;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 1 ATGAAAGGATCTTAATCCATTTTATCTCTTCGATTCCGTTAACCCCGCAATCTGCA 60
DB 1 ATGAAAGGATCTTAATCCATTTTATCTCTTCGATTCCGTTAACCCCGCAATCTGCA 60
QY 61 TTCGCTCAGAGTGAGCCGAGAGCTGAAGCTGGAAGTGTGATTTGATGTCATGATGT 120
DB 61 TTCCTCAGAGTGAGCCGAGAGCTGAAGCTGGAAGTGTGATTTGATGTCATGATGT 120
QY 121 GTGGGTGCTCAACCAAGGCGACGCAACTGATGAGGATGTCACCCGAGAGCATGAGCA 180
DB 121 GTGGGTGCTCAACCAAGGCGACGCAACTGATGAGGATGTCACCCGAGAGCATGAGCA 180
QY 181 ACCTGCGCGTAAACTGAGTGAAGCTGACACCGCGCGGTGATGCTAATCTC 240
DB 181 ACCTGCGCGTAAACTGAGTGAAGCTGACACCGCGCGGTGATGCTAATCTC 240
QY 241 GAGCATTAATGCGCTCAGCGTCTGATGAGCGACGAGATTGCTGCTAAATGTGCTGCCG 300
DB 241 GAGCATTAATGCGCTCAGCGTCTGATGAGCGACGAGATTGCTGCTAAATGTGCTGCCG 300
QY 301 CAGTCTGCTCAGGTGCGGATTAATGCTGATGTCACGACGCTAACCCGTTAAACAGGCGAA 360
DB 301 CAGTCTGCTCAGGTGCGGATTAATGCTGATGTCACGACGCTAACCCGTTAAACAGGCGAA 360
QY 361 GCCTTCGCGCGCGGCTGAGCACTGACTGTGCAATTAACCTAATACCCAGGAGATACG 420
DB 361 GCCTTCGCGCGCGGCTGAGCACTGACTGTGCAATTAACCTAATACCCAGGAGATACG 420

DB 361 GCCTTCGCGCGCGGCTGAGCACTGACTGTGCAATTAACCTAATACCCAGGAGATACG 420
QY 421 TCCAGTCCCGATCCGTTAATTTAATCCCTTAATAAATCGGCGCTTGCCAACTGTAATACGG 480
DB 421 TCCAGTCCCGATCCGTTAATTTAATCCCTTAATAAATCGGCGCTTGCCAACTGTAATACGG 480
QY 481 AACGTGACTGACGATCTCTGAGAGGGCAGAGGGTCAATTGCTGACTTTACCGGGCAT 540
DB 481 AACGTGACTGACGATCTCTGAGAGGGCAGAGGGTCAATTGCTGACTTTACCGGGCAT 540
QY 541 TATCAACGCGCTTTTCGCAACTGGAACGCGGTGCTTAATTTTCGCAATCAACTGTGC 600
DB 541 CGGCAAAACGCGCTTTTCGCAACTGGAACGCGGTGCTTAATTTTCGCAATCAACTGTGC 600
QY 601 CTTAAAGTGAAGAAACAGACGAAAGCTGTTCAATTAAACGAGCATTAACATGGAATC 660
DB 601 CTTAAAGTGAAGAAACAGACGAAAGCTGTTCAATTAAACGAGCATTAACATGGAATC 660
QY 661 AAGGTGAGCGCGACTGTGCTCATTTAACCGGTGCGGTAAAGCTGCAATGATGTCAG 720
DB 661 AAGGTGAGCGCGACTGTGCTCATTTAACCGGTGCGGTAAAGCTGCAATGATGTCAG 720
QY 721 GAGATATTTCTCTGCAACAGCAAGGAATGCCGAGCCGCGGTGGAAGATCAC 780
DB 721 GAGATATTTCTCTGCAACAGCAAGGAATGCCGAGCCGCGGTGGAAGATCAC 780
QY 781 GATTCAACAGATGGAACACCTTGCTTAATGTTGATTAACCGGCAATTTGATTTGCTTAA 840
DB 781 GATTCAACAGATGGAACACCTTGCTTAATGTTGATTAACCGGCAATTTGATTTGCTTAA 840
QY 841 CGCAGCGCAGAGGTGGCGCGCAGCGCGCAGCCCGTATTATTAATTTGATCAAGACGCG 900
DB 841 CGCAGCGCAGAGGTGGCGCGCAGCGCGCAGCCCGTATTATTAATTTGATCAAGACGCG 900
QY 901 TTGACGCCCATCCACCGCAAAAAACAGCGGTATGTGATGATTAACCATTCATGCTG 960
DB 901 TTGACGCCCATCCACCGCAAAAAACAGCGGTATGTGATGATTAACCATTCATGCTG 960
QY 961 TTTATGCCCGGACACGATTAATCTGAGCAAACTCGGCGCGCAGCTGAGCTCAACTGG 1020
DB 961 TTTATGCCCGGACACGATTAATCTGAGCAAACTCGGCGCGCAGCTGAGCTCAACTGG 1020
QY 1021 ACGCTTCGCGGTGAGCGGATTAACAGCGCGCAGGTGATGATGCTGTTGAACGCTGG 1080
DB 1021 ACGCTTCGCGGTGAGCGGATTAACAGCGCGCAGGTGATGATGCTGTTGAACGCTGG 1080
QY 1081 CGTGGCTAAGCATTAACAGCGAGTGTCAAGTTCGCTGTGCTTTCAGACTTTACAG 1140
DB 1081 CGTGGCTAAGCATTAACAGCGAGTGTCAAGTTCGCTGTGCTTTCAGACTTTACAG 1140
QY 1141 CAGATGCTGATTAACAGCGCGCTGTCAATTAATGCGCGCGGAGAGGTGAACAGTACC 1200
DB 1141 CAGATGCTGATTAACAGCGCGCTGTCAATTAATGCGCGCGGAGAGGTGAACAGTACC 1200
QY 1201 CTGGCAGATGTGAAGAGCGAAATGCGCAGGAGATGTTCTGTTGCGAGGTTTACGCA 1260
DB 1201 CTGGCAGATGTGAAGAGCGAAATGCGCAGGAGATGTTCTGTTGCGAGGTTTACGCA 1260
QY 1261 ATGTGAATGAAGCAGCATACCGCGCTGACGTTTGAATCTCATC 1306
DB 1261 ATGTGAATGAAGCAGCATACCGCGCTGACGTTTGAATCTCATC 1306
RESULT 10
AAD36473
ID AAD36473 standard; DNA; 1323 BP.
XX
XX AAD36473;
XX
XX 21-AUG-2002 (first entry)
XX
XX Escherichia coli phytase DNA.
XX

DE Escherichia coli B phytase DNA.
XX
XX Bacterial phytase; K12 appa phytase; protease stability; anabolic;
KM gastrointestinal; nutritional value; feed treatment process; therapy;
KM thermal tolerance; growth performance; alcoholic drink; biopulping;
KM non-alcoholic drink; bioleaching; B phytase; ds.
XX
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT CDS 1..1323
FT /tag= a
FT /product= "E. coli B phytase protein"
XX
XX WO200190333-A2.
XX
XX 29-NOV-2001.
XX
XX 24-MAY-2001; 2001WO-US017118.
XX
XX 25-MAY-2000; 2000US-00580515.
XX
XX (DIVE-) DIVERSA CORP.
XX
XX Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI WPI; 2002-083108/11.
XX P-PSDB; AAE15806.
XX
XX New bacterial phytase for e.g. improving the nutritional value of phytate
PT -containing foodstuffs and subsequently improving the growth performance
PT of an organism that consumes it, or in treating animal digestive systems.
XX
XX Claim 2; Fig 1; 170pp; English.
XX
XX The patent discloses recombinant bacterial phytase from Escherichia coli
CC K12 appa phytase. The enzyme has phytase activity and improved thermal
CC tolerance when compared with wild-type phytase. It has improved protease
CC stability at low pH. The recombinant phytase is useful for improving the
CC nutritional value of phytate-containing foodstuffs and subsequently
CC improving the growth performance of an organism that consumes it, in
CC treating animal digestive systems, in feed treatment processes and for in
CC vitro purposes related to research, discovery and development. They are
CC also used for generating recombinant digestive system life forms, for
CC producing or manufacturing alcoholic and non-alcoholic drinks based on
CC the use of moulds, grains and/or plants, in biopulping and bio-bleaching
CC where a reduction in the use of environmentally harmful chemicals that
CC are traditionally used in the pulp and paper industry is desired and in
CC the reduction or possible elimination of the need for mineral
CC supplements, enzymes or therapeutic drugs for animals from the daily feed
CC thus increasing the amount calories and nutrients present in the feed.
CC The present sequence is a DNA encoding E. coli B phytase protein
XX
SQ Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;
Query Match 96.7%; Score 1265; DB 6; Length 1323;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 1 ATGAAGGATCTTAATCCATTTTATCTTCTGATTCGGTTAACCCCGCAATCTGCA 60
DB 1 ATGAAGGATCTTAATCCATTTTATCTTCTGATTCGGTTAACCCCGCAATCTGCA 60
QY 61 TTCGCTCAGAGTAGCCGAGAGCTGAAGCTGAAGTGTGTGATTTGCAATGTCATGCT 120
DB 61 TTCCTCTCAGAGTAGCCGAGAGCTGAAGTGTGTGATTTGCAATGTCATGCT 120
QY 121 GTGGGTGCTCAACCAAGGCGACGCACTGATGCAAGATGCAACCCGACGAGCATGGCA 180
DB 121 GTGGGTGCTCAACCAAGGCGACGCACTGATGCAAGATGCAACCCGACGAGCATGGCA 180
QY 181 ACCTGCGCGGTAAACTGGGTAGCTGACACCGCGCGGTGGTGAAGTAAATGCGCTATCTC 240
DB 181 ACCTGCGCGGTAAACTGGGTAGCTGACACCGCGCGGTGGTGAAGTAAATGCGCTATCTC 240

DB 181 ACCTGCGCGGTAAACTGGGTAGCTGACACCGCGCGGTGGTGAAGTAAATGCGCTATCTC 240
QY 241 GGAATTAATCTGGGGTCAAGCGCTCTGGTATGCCGAGGATTTGCTTAATATGTGGTCCG 300
DB 241 GGAATTAATCAAGCGCGCTCTGGTATGCCGAGGATTTGCTGGGAAAAAGGGCTGCGCG 300
QY 301 CAGTCTGGTCAAGGTCGGAATTAATGCTGATGTGACAGAGGTCACCGGTAAACAGGCGAA 360
DB 301 CAGTCTGGTCAAGGTCGGAATTAATGCTGATGTGACAGAGGTCACCGGTAAACAGGCGAA 360
QY 361 GCCTTCGCGCGCGGGCTGGGACCTGACTGTGCAATTAACCGTACATCAACGAGCATACG 420
DB 361 GCCTTCGCGCGCGGGCTGGGACCTGACTGTGCAATTAACCGTACATCAACGAGCATACG 420
QY 421 TCCAGTCCGATCCGTTATTTAATCCCTTAATAAACTGGCGGTTGGCAACGGAATAAGCG 480
DB 421 TCCAGTCCGATCCGTTATTTAATCCCTTAATAAACTGGCGGTTGGCAACGGAATAAGCG 480
QY 481 AACGTACTGACGCGATCTCTGAGAGGGCAGAGGGTCAATTTGCTGACTTTACCGGCGAT 540
DB 481 AACGTACTGACGCGATCTCTGAGAGGGCAGAGGGTCAATTTGCTGACTTTACCGGCGAT 540
QY 541 TATCAAAACGCGGTTTCGCGAATGGAACGAGTGTCTTAATTTTCGCAATCAAACTGTGC 600
DB 541 TATCAAAACGCGGTTTCGCGAATGGAACGAGTGTCTTAATTTTCGCAATCAAACTGTGC 600
QY 601 CTTTAAAGTGAAGAAACGAGACGAAAGCTGTTCATTAACGAGCATTAACATCGGAATC 660
DB 601 CTTTAAAGTGAAGAAACGAGACGAAAGCTGTTCATTAACGAGCATTAACATCGGAATC 660
QY 661 AAGGTAGCCCGCGACTGTCTCATTTAAACGAGTGGGTGAAGCTCGCATCAATGCTGACG 720
DB 661 AAGGTAGCCCGCGACTGTCTCATTTAAACGAGTGGGTGAAGCTCGCATCAATGCTGACG 720
QY 721 GAGATATTTCTCTGCAACAGCAAGGAAATCCGAGCCGAGGTGGGAAAGATCAAC 780
DB 721 GAGATATTTCTCTGCAACAGCAAGGAAATCCGAGCCGAGGTGGGAAAGATCAAC 780
QY 781 GATTCAACACAGTGGAAACACTTGTCTAAGTTTGATTAACCGCAATTTGATTTGCTCA 840
DB 781 GATTCAACACAGTGGAAACACTTGTCTAAGTTTGATTAACCGCAATTTGATTTGCTCA 840
QY 841 CGACGCGCAGAGGTTTCCGCAACAGCAAGGAAATCCGAGCCGAGGTGGGAAAGATCAAC 900
DB 841 CGACGCGCAGAGGTTTCCGCAACAGCAAGGAAATCCGAGCCGAGGTGGGAAAGATCAAC 900
QY 901 TTGACGCCCATCCACCGCAAAAACAGGCGTATGTGTGACATTAACCATTCAGTACTG 960
DB 901 TTGACGCCCATCCACCGCAAAAACAGGCGTATGTGTGACATTAACCATTCAGTACTG 960
QY 961 TTTATGCGCGGACACGATTAATCTTGTGGCAAAATCTCGCGCGGACATGAGCTCAACTG 1020
DB 961 TTTATGCGCGGACACGATTAATCTTGTGGCAAAATCTCGCGCGGACATGAGCTCAACTG 1020
QY 1021 ACGCTTCCCGGTGACCGGATTAACAGCCCGCAGAGTGTGAATGCTGTTGAACGCTGG 1080
DB 1021 ACGCTTCCCGGTGACCGGATTAACAGCCCGCAGAGTGTGAATGCTGTTGAACGCTGG 1080
QY 1081 CGTGGGTTAAGCGATTAACAGCCAGGTGATTCAGGTTTCGCTGTCTTTCAGACTTTACG 1140
DB 1081 CGTGGGTTAAGCGATTAACAGCCAGGTGATTCAGGTTTCGCTGTCTTTCAGACTTTACG 1140
QY 1141 CAGATGCTGATTAATAACCGCGCTGTCTTAATATCGCGCCCGGAGAGGTGAACCTGACC 1200
DB 1141 CAGATGCTGATTAATAACCGCGCTGTCTTAATATCGCGCCCGGAGAGGTGAACCTGACC 1200
QY 1201 CTGGCAGATGTGAAGAGCGAAATGCGCAGAGGCAATGTTCGTTGGCAGAGTTTACGCAA 1260
DB 1201 CTGGCAGATGTGAAGAGCGAAATGCGCAGAGGCAATGTTCGTTGGCAGAGTTTACGCAA 1260
QY 1261 ATCGTGAATGAAGCAGCATACCGGCGTGCAGTTTGAATCTCATC 1306
DB 1261 ATCGTGAATGAAGCAGCATACCGGCGTGCAGTTTGAATCTCATC 1306

RESULT 12
ADA19445
ID ADA19445 standard, DNA, 1323 BP.
XX
XX ADA19445;
XX
XX 20-NOV-2003 (first entry)
XX
XX E. coli B DNA encoding phytase.
XX
XX Phytase; ds; gene; phytase; appa gene; animal feed; inorganic phosphate;
XX digestion enhancement; transgenic; thermal tolerance; protease stability.
OS Escherichia coli B.
XX
XX
XX Key Location/Qualifiers
XX CDS 1..1323
XX FT /*tag= a
XX FT /product= "Phytase"
XX
XX US2002136754-A1.
XX
XX 26-SEP-2002.
XX
XX 24-MAY-2001; 2001US-00866379.
XX
XX 13-AUG-1997; 97US-00910798.
XX 01-MAR-1999; 99US-00259214.
XX 13-APR-1999; 99US-00291931.
XX 25-MAY-1999; 99US-00318528.
XX 25-MAY-2000; 2000US-00580515.
XX
XX (SHOR/) SHORT J M.
XX (KRET/) KRETTZ K A.
XX (GRAY/) GRAY K A.
XX (BART/) BARTON N R.
XX (GARR/) GARRETT J B.
XX (DONO/) O' DONOGHUE E.
XX (MATH/) MATHUR E J.
XX
XX Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB;
XX O' Donoghue E, Mathur EJ;
XX
XX WPI: 2003-040002/03.
XX P-PSDB; ADA19446.
XX
XX Isolated Escherichia coli polymucleotide encoding a modified phytase
XX enzyme, useful in the production of animal feed, for improving the
XX nutritional value of phytase-containing feedstuff and for enhancing
XX digestion in humans and animals.
XX
XX Claim 2; Fig 1; 62pp; English.
XX
XX The invention relates to an isolated Escherichia coli polymucleotide
XX encoding a phytase enzyme appearing as ADA19450 and having amino acids
XX modification from W68E, Q84W, A95P, K97C, S168E, R180Y, N26C or Y277D.
XX Also included the E. coli appa gene ADA19449 (or an oligonucleotide
XX derived from it) or its mutant sequence ADA19452, expression vectors,
XX host cells, a method of improving nutritional value of a phytase-
XX containing feedstuff by contacting the phytase-containing feedstuff with
XX a pure phytase enzyme mutated phytase (where the phytase enzyme catalyzes
XX the liberation of inorganic phosphate from the phytase in the phytase-
XX containing feedstuff), a method to produce an animal feed containing a
XX microbial phytase (comprising culturing the plant cell, plant part or
XX plant under conditions where the nucleotide sequence is expressed and
XX converting the plant cells, plants or plant into a composition for animal
XX feed), a feed composition for animals (comprising the plant seeds, plant
XX cells, plant parts or plants in admixture with a phytase-containing
XX feedstuff), a method to treat a human or an animal able to benefit from
XX digestive enhancement by the activity of an exogenous phytase enzymes
XX comprising administering to the human or animal the plant seed, plant

CC cells, plant parts or plants of a transgenic plant which is modified to
CC contain an expression system which expresses a nucleotide sequence
CC encoding a phytase enzyme, a transgenic non-human organism whose genome
CC comprising a heterologous nucleic acid sequence encoding a polypeptide
CC having phytase activity. The phytase enzyme is useful for improving the
CC nutritional value of phytase-containing feedstuff, in the production of
CC animal feed and for enhancing digestion in humans and animals. The
CC invented method improves thermal tolerance and protease stability. It
CC also improves the feeding value of phytase rich ingredients. The present
CC sequence encodes E. coli B wild-type phytase.
XX
XX
SQ Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;
Query Match 96.7%; Score 1265; DB 8; Length 1323;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 1 ATGAAGGAGATCTTAATCCATTTTATCTCTGATTCGGTTAAACCCCGAATCTGCA 60
DB 1 ATGAAGGAGATCTTAATCCATTTTATCTCTGATTCGGTTAAACCCCGAATCTGCA 60
QY 61 TTGCTCAGAGTGAGCCGGAAGCTGAAAGTGTGTGATTTGTGATGTCATGAT 120
DB 61 TTGCTCAGAGTGAGCCGGAAGCTGAAAGTGTGTGATTTGTGATGTCATGATGAT 120
QY 121 GTGCTGCTCCAAACCAAGCCCAAGCAATGATGCAAGATGTCACCCAGACGATGCGCA 180
DB 121 GTGCTGCTCCAAACCAAGCCCAAGCAATGATGCAAGATGTCACCCAGACGATGCGCA 180
QY 181 ACTTGACCGGTAATACTGGGTGAGCTGACACCGCGGGTGGTGAAGCTATGCTATGTC 240
DB 181 ACTTGACCGGTAATACTGGGTGAGCTGACACCGCGGGTGGTGAAGCTATGCTATGTC 240
QY 241 GGCATTTACTGAGCGTCAAGCTGTGGTACCGACGAGATTGCTCTTAAATGTGTGCTCCG 300
DB 241 GGCATTTACTGAGCGTCAAGCTGTGGTACCGACGAGATTGCTCTTAAATGTGTGCTCCG 300
QY 301 CAGTCTGCTCAGTGTGCGATTTATGCTGATGTCGACGAGCGTAAACCGGTAACAGCGCA 360
DB 301 CAGTCTGCTCAGTGTGCGATTTATGCTGATGTCGACGAGCGTAAACCGGTAACAGCGCA 360
QY 361 GCGTTGCGCGCGCGGCTGGGCACTGACCTGATCTCAATTAACCGTACATCCAGGCAATACG 420
DB 361 GCGTTGCGCGCGGCTGGGCACTGACCTGATCTCAATTAACCGTACATCCAGGCAATACG 420
QY 421 TCCAGTCCCGATCCGTTATTTATCTCTTAATAAACTGCGCTTTGCAACTGGAATACCG 480
DB 421 TCCAGTCCCGATCCGTTATTTATCTCTTAATAAACTGCGCTTTGCAACTGGAATACCG 480
QY 481 AACGTGACTGACGCGATCTCTGAGAGGCGAGAGGCTCAATGCTGACTTTAACCGGGCAT 540
DB 481 AACGTGACTGACGCGATCTCTGAGAGGCGAGAGGCTCAATGCTGACTTTAACCGGGCAT 540
QY 541 TATCAACGCGGCTTTCGGAACGGAACGGTCTTAATTTCCGCAATCAAACTTGTGC 600
DB 541 TATCAACGCGGCTTTCGGAACGGAACGGTCTTAATTTCCGCAATCAAACTTGTGC 600
QY 601 CTTAAACGTGAGAAACAGAACGAAAGCTGTTCAATTAACGACGAGATTAACATCGGAATC 660
DB 601 CTTAAACGTGAGAAACAGAACGAAAGCTGTTCAATTAACGACGAGATTAACATCGGAATC 660
QY 661 AAGGTGAGCGCGCATGTGTCTCAATTAACCGGTGCGGTAAAGCTTGCATCAATCTGACG 720
DB 661 AAGGTGAGCGCGCATGTGTCTCAATTAACCGGTGCGGTAAAGCTTGCATCAATCTGACG 720
QY 721 GAGATATTTTCTCGCAACAGCAAGCAAGATGCGGAGCGGGGTGGGAAGATCAAC 780
DB 721 GAGATATTTTCTCGCAACAGCAAGCAAGATGCGGAGCGGGGTGGGAAGATCAAC 780
QY 781 GATTTCACACGAGTGAACACCTTGCTAAAGTTTGCAATACGCGCAATTTGATTTGCTACAA 840
DB 781 GATTTCACACGAGTGAACACCTTGCTAAAGTTTGCAATACGCGCAATTTTATTGCTACAA 840

```
Qy 841 CGCAGCCAGAGTGTGCCCCGAGCGCGCCACCCTGTTATTAAGATTGATCAAGACAGCG 900
Db 841 CGCAGCCAGAGTGTGCCCCGAGCGCGCCACCCTGTTATTAAGATTGATCAAGACAGCG 900
Qy 901 TTGACGCCCATCCACCGCAAAAAACAGCGGTATGTGTGACATTACCACTTCAGTCTG 960
Db 901 TTGACGCCCATCCACCGCAAAAAACAGCGGTATGTGTGACATTACCACTTCAGTCTG 960
Qy 961 TTTATGCCCGGACACGATTAATCTGCAAAATCCGCGCGCGGACCTGAGAGCTCAACTCG 1020
Db 961 TTTATGCCCGGACACGATTAATCTGCAAAATCTGCGCGCGGACCTGAGAGCTCAACTCG 1020
Qy 1021 ACGCTTCCCGGTCAACCGGATATACACGCGCCAGGTGTGAACTGTGTGTTGAACGCTG 1080
Db 1021 ACGCTTCCCGGTCAACCGGATATACACGCGCCAGGTGTGAACTGTGTGTTGAACGCTG 1080
Qy 1081 CGTCGGCTAAGCGATTAACAGCAGCTGATTCAGGTTTCGTGCTTCACAGCTTTACAG 1140
Db 1081 CGTCGGCTAAGCGATTAACAGCAGCTGATTCAGGTTTCGTGCTTCACAGCTTTACAG 1140
Qy 1141 CAGATGGGTGATTAAGCGCGCTGTCAATTAATAGCGCGCGCGGAGAGGTGAACCTGACC 1200
Db 1141 CAGATGGGTGATTAAGCGCGCTGTCAATTAATAGCGCGCGGAGAGGTGAACCTGACC 1200
Qy 1201 CTGGCAGGATGTGAAGACGCAAAATGCGCAGGCGCATGTGTGCTGGCAGGTTTACGCAA 1260
Db 1201 CTGGCAGGATGTGAAGACGCAAAATGCGCAGGCGCATGTGTGCTGGCAGGTTTACGCAA 1260
Qy 1261 ATCGTGAATGAAGACGCGATACCGGCGGTGACGTTTGAGANTCTCATC 1306
Db 1261 ATCGTGAATGAAGACGCGATACCGGCGGTGAGATCTCATC 1306
```

RESULT 13

```
AD050291
ID AD050291 standard; DNA; 1323 BP.
AC AD050291;
XX
DT 29-JUN-2004 (first entry)
DE Escherichia coli B phytase DNA.
XX
KM Phytase; bacteria; thermal tolerance; protease stability; foodstuff;
KW phytase; animal feed; fish feed; dough; baking; gene; ds.
XX
OS Escherichia coli; B.
XX
FH Key Location/Qualifiers
FT CDS 1..1323
FT /tag= a
FT /product= "Phytase enzyme"
FT /tag= b
FT /note= "Encodes Arg"
XX
PN US2004091968-A1.
XX
PD 13-MAY-2004.
XX
PF 20-JUN-2003; 2003US-00601319.
XX
PR 13-AUG-1997; 97US-00910798.
PR 01-MAR-1999; 99US-00259214.
PR 13-APR-1999; 99US-00291931.
PR 25-MAY-1999; 99US-00318528.
PR 25-MAY-2000; 2000US-00580515.
PR 24-MAY-2001; 2001US-00866379.
XX
PA (SHOR/) SHORT J M.
PA (KRET/) KRETZ K.
PA (GRAY/) GRAY K A.
PA (BART/) BARTON N R.
```

```
PA (GAR/) GARRETT J B.
PA (ODON/) O'DONOGHUE B.
PA (MATH/) MATHER E J.
PI Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI Mather EJ;
DR WPI: 2004-374952/35.
DR P-PDSB: AD050292.
XX
XX Producing phytase, involves providing nucleic acid derived from bacteria
PT encoding polypeptide having phytase activity, and expressing nucleic acid
PS in yeast.
PS Claim 3; SEQ ID NO 1; 74pp; English.
XX
XX The invention relates to a method for producing phytase that involves
CC providing a nucleic acid encoding phytase derived from a bacteria, and
CC expressing the nucleic acid in a yeast under conditions that allow
CC expression of the enzyme in the yeast. The invention also relates to
CC modified phytase enzyme which has improved thermal tolerance and protease
CC stability at low pH. The phytase enzyme can be used in foodstuffs to
CC improve the feeding value of phytase rich ingredients, and in diet of
CC numerous animals including mammals, fowls and fishes, commercially
CC significant mammals such as pigs, goats, laboratory rodents, commercially
CC significant avian species such as chicken, ducks, doves, parrot, etc.,
CC commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC etc., in dough making and baking, in dietary aids for animals. The method
CC provides easy manufacture of the active ingredient loaded biocompatible
CC composition, higher yields and loading efficiency. The phytase
CC incorporated in to the dietary aid is safe for animals. The present
CC sequence is Escherichia coli B phytase DNA.
XX
SQ Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;
```

```
Query Match 96.7%; Score 1265; DB 12; Length 1323;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
```

```
Qy 1 ATGAAAGCGATCTTAATCCATTTTATCTCTCTGATTCGCTTAACCCCGCAATCGCA 60
Db 1 ATGAAAGCGATCTTAATCCATTTTATCTCTCTGATTCGCTTAACCCCGCAATCGCA 60
Qy 61 TTGCGTCAAGTGAAGCCGAGCTGAAGCTGAAAGTGTGATTCAGTCCGTCATGT 120
Db 61 TTGCGTCAAGTGAAGCCGAGCTGAAGCTGAAAGTGTGATTCAGTCCGTCATGT 120
Qy 121 GTGCGTCTCCAAACCAAGGCCACGCACTGATGCAAGATGCAACCCAGACGATGGCA 180
Db 121 GTGCGTCTCCAAACCAAGGCCACGCACTGATGCAAGATGCAACCCAGACGATGGCA 180
Qy 181 ACCTGGCCGGTAAAACTGGGTGAGCTGACACCGCGCGGTGTGAGCTAAATCGCTATCTC 240
Db 181 ACCTGGCCGGTAAAACTGGGTGAGCTGACACCGCGGAGTGTGAGCTAAATCGCTATCTC 240
Qy 241 GGAATTACTGGCGTCAACGCTGTGTGAGCCGAGATTTGCTGCTAAATGTGCTGCGCG 300
Db 241 GGAATTACTGGCGTCAACGCTGTGTGAGCCGAGATTTGCTGCGGAAAGGCGTGC 300
Qy 301 CAGTGTGTGAGGTGCGGATTAATTTGCTGATGTGACAGAGGTACCGGTAAACAGGGCAA 360
Db 301 CAGTGTGTGAGGTGCGGATTAATTTGCTGATGTGACAGAGGTACCGGTAAACAGGGCAA 360
Qy 361 GCGTTGCGCGCGCGGCGTGGACCTGATGTGCAATAACCTATCAATACCCAGGAGATACG 420
Db 361 GCGTTGCGCGCGCGGCGTGGACCTGATGTGCAATAACCTATCAATACCCAGGAGATACG 420
Qy 421 TCCAGTCCGATCGGTTATTTAATCCTCTAAATACTGCGCTTGGCCACTGATGATACGGC 480
Db 421 TCCAGTCCGATCGGTTATTTAATCCTCTAAATACTGCGCTTGGCCACTGATGATACGGC 480
Qy 481 AACGTGCTGACGCGATCTCGAAGGGCAGAGGGTCAAATGCTGACTTACCGGGCAT 540
Db 481 AACGTGCTGACGCGATCTCGAAGGGCAGAGGGTCAAATGCTGACTTACCGGGCAT 540
```

```

Db 481 AACGTAAGTGAACGCAATCTCTCAGACGAGGAGGAGGTCAATGCTGATCTTAACCGGCGAT 540
Qy 541 TATCAACGCGGCTTTCGCGACTGGAACGGGTGCTTAATTTCCGGAATCAACCTTGTGC 600
Db 541 CGGCAACGCGGCTTTCGCGACTGGAACGGGTGCTTAATTTCCGGAATCAACCTTGTGC 600
Qy 601 CTTAAACGTGAGAAACAGACGAAAGCTGTCTTAACGAGGATTAACATCGAATC 660
Db 601 CTTAAACGTGAGAAACAGACGAAAGCTGTCTTAACGAGGATTAACATCGAATC 660
Qy 661 AAGGTAGCGCCGAGCTGTCTCTTAACCGGTGCGTAAAGCTCGCATCAATGCTGAGC 720
Db 661 AAGGTAGCGCCGAGCTGTCTCTTAACCGGTGCGTAAAGCTCGCATCAATGCTGAGC 720
Qy 721 GAGATATTTCTCTGGAACAAGCAACGGAATGCGGAGCCGGGTGGGAAAGATCAAC 780
Db 721 GAGATATTTCTCTGGAACAAGCAACGGAATGCGGAGCCGGGTGGGAAAGATCAAC 780
Qy 781 GATTCAACACGATGGAACAACCTTGCTAAGTTGCAATTAACGCGCAATTTGATTTGCTACA 840
Db 781 GATTCAACACGATGGAACAACCTTGCTAAGTTGCAATTAACGCGCAATTTGATTTGCTACA 840
Qy 841 CGCAGCGCAAGAGTTGCGCGCAGCCGCGCACCCGTTAATTAAGATTGATCAAGACGCG 900
Db 841 CGCAGCGCAAGAGTTGCGCGCAGCCGCGCACCCGTTAATTAAGATTGATCAAGACGCG 900
Qy 901 TTGAGCGCCCATCCACCGGAAACAGGGGTATGCTGACATTACCACTTCACTGCTG 960
Db 901 TTGAGCGCCCATCCACCGGAAACAGGGGTATGCTGACATTACCACTTCACTGCTG 960
Qy 961 TTTATCGCGGACAGATTAATCTATGCGCAATCTCGGCGGCGCATCGGAGCTCAACTG 1020
Db 961 TTTATCGCGGACAGATTAATCTATGCGCAATCTCGGCGGCGCATCGGAGCTCAACTG 1020
Qy 1021 ACGCTTCCCGGTGACCGCGATTAACACGCGCGCAGGTGCTGAACTGATTTGAAACGCTG 1080
Db 1021 ACGCTTCCCGGTGACCGCGATTAACACGCGCGCAGGTGCTGAACTGATTTGAAACGCTG 1080
Qy 1081 CGTGGCGTAAAGGATTAACGCGGATGATTCAGGTTTGCTGCTTCCAGACTTTACG 1140
Db 1081 CGTGGCGTAAAGGATTAACGCGGATGATTCAGGTTTGCTGCTTCCAGACTTTACG 1140
Qy 1141 CAGATCGCTGATTAACACGCGCTGCTCATTAATAACGCGCGCGGAGAGGTGAACCTGACC 1200
Db 1141 CAGATCGCTGATTAACACGCGCTGCTCATTAATAACGCGCGCGGAGAGGTGAACCTGACC 1200
Qy 1201 CTGGCAGGATGTAAGAGCGAATGCGACGCGCATGTGTGCGGAGTTTACGCA 1260
Db 1201 CTGGCAGGATGTAAGAGCGAATGCGACGCGCATGTGTGCGGAGTTTACGCA 1260
Qy 1261 ATCGGATGTAAGAGCGAATGCGCGGCTGCTGATTAATAACGCGCGGAGAGGTGAACCTGACC 1306
Db 1261 ATCGGATGTAAGAGCGAATGCGCGGCTGCTGATTAATAACGCGCGGAGAGGTGAACCTGACC 1306

RESULT 14
AD050301 standard; DNA, 1323 BP.
XX AD050301;
XX 29-JUN-2004 (first entry)
XX Kangaroo rat Escherichia coli phytase 872PHI mutant DNA.
XX Phytase; bacteria; thermal tolerance; protease stability; foodstuff;
XX phytase; animal feed; fish feed; dough; baking; mutant; gene; de.
XX Escherichia coli.
XX Key Location/Qualifiers
XX CDS 1..1323
XX FT /*tag= a

```

```

FT /product= "Phytase mutant enzyme"
FT 214..216
FT unsure
FT /*tag= b
FT /note= "Encodes Arg"
FT replace(527,A)
FT mutation
FT /*tag= c
FT replace(893,T)
FT mutation
FT /*tag= d
FT replace(895,G)
FT mutation
FT /*tag= e
FT replace(934,G)
FT /*tag= f
PN US2004091968-A1.
XX 13-MAY-2004.
PD 20-JUN-2003; 2003US-00601319.
XX 13-AUG-1997; 97US-00910798.
PR 01-MAR-1999; 99US-00259214.
PR 13-APR-1999; 99US-00291931.
PR 25-MAY-1999; 99US-00318528.
PR 25-MAY-2000; 2000US-00580515.
PR 24-MAY-2001; 2001US-00866379.
XX (SHORT) SHORT J M.
XX (KRETT) KRETT K.
XX (GRAY) GRAY K A.
XX (BART) BARTON N R.
XX (GARR) GARRATT J B.
XX (ODON) O'DONOGHUE E.
XX (MATH) MATHER E J.
PI Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI Mather EJ;
PI WPI: 2004-374952/35.
PI P-PSDB; AD050302.
XX Producing phytase, involves providing nucleic acid derived from bacteria
XX encoding polypeptide having phytase activity, and expressing nucleic acid
XX in yeast.
XX Disclosure; Page; 74pp; English.
XX The invention relates to a method for producing phytase that involves
XX providing a nucleic acid encoding phytase derived from a bacteria, and
XX expressing the nucleic acid in a yeast under conditions that allow
XX expression of the enzyme in the yeast. The invention also relates to
XX modified phytase enzyme which has improved thermal tolerance and protease
XX stability at low pH. The phytase enzyme can be used in foodstuffs to
XX improve the feeding value of phytase rich ingredients, and in diet of
XX numerous animals including mammals, fowls and fishes, commercially
XX significant avian species such as pigs, goats, laboratory rodents, commercially
XX commercially farmed fish such as guppy, zebrafish, molly, swordtail,
XX etc., in dough making and baking, in dietary aids for animals. The method
XX provides easy manufacture of the active ingredient loaded biocompatible
XX composition, higher yields and loading efficiency. The phytase
XX incorporated in to the dietary aid is safe for animals. The present
XX sequence is Kangaroo rat Escherichia coli phytase mutant DNA. Note: The
XX present sequence is not shown in the specification but has been derived
XX from Escherichia coli B phytase DNA AD050291.
XX Sequence 1323 BP; 325 A; 353 C; 356 G; 288 T; 0 U; 1 Other;
XX Query Match 96.7%; Score 1265; DB 12; Length 1323;
XX Best Local Similarity 98.0%; Pred. No. 0;
XX Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
Qy 1 ATGAAGGATCTTAATCCGATTTTATCTCTGATTCGTTACCGCGCATTCGCA 60

```

```

Db      1 ATGAAACGATCTTAATCCATTTTATCTCTGATTCCTGTAACCCCGCAATCTGCA 60
Qy      61 TTGCTCAGAGTGAAGCCGAGCTGAGCTGGAAGTGTGATTTGTCAGTCCGTCATGCT 120
Db      61 TTCCCTCAGAGTGAAGCCGAGCTGAGCTGGAAGTGTGATTTGTCAGTCCGTCATGCT 120
Qy      121 GTGCGTCTCCAAACCAAGGCCACCACTGATGCAAGATTCACCCGAGCGCATGCGCA 180
Db      121 GTGCGTCTCCAAACCAAGGCCACCACTGATGCAAGATTCACCCGAGCGCATGCGCA 180
Qy      181 ACCTGCGCGGTAAACCTGAGTGAAGTGAACCGCGCGGTGTGAGCTAATGCTATCTC 240
Db      181 ACCTGCGCGGTAAACCTGAGTGAAGTGAACCGCGCGGTGTGAGCTAATGCTATCTC 240
Qy      241 GGAATTAATCTGCGTCAAGCGCTGCTGATGCGGACGGAATGCTCTAAATGAGCGCGCG 300
Db      241 GGAATTAATCTGCGTCAAGCGCTGCTGATGCGGACGGAATGCTCTAAATGAGCGCGCG 300
Qy      301 CAGTCTGATCAGTCCGATTAATGCTGATGTCAGACGACGCTACCCGTAACAGCGGCA 360
Db      301 CAGTCTGATCAGTCCGATTAATGCTGATGTCAGACGACGCTACCCGTAACAGCGGCA 360
Qy      361 GCTTCGCGCGCGCGCTGCGCACTGATCTGCAATTAACCGTAATCCAGCGCATACG 420
Db      361 GCTTCGCGCGCGCGCTGCGCACTGATCTGCAATTAACCGTAATCCAGCGCATACG 420
Qy      421 TCCAGTCCCGATCGGTTATTAATCCTTAATAACCTGCGCTTGCACATGAGATACGCG 480
Db      421 TCCAGTCCCGATCGGTTATTAATCCTTAATAACCTGCGCTTGCACATGAGATACGCG 480
Qy      481 AACGTGACTGACGCGATCCTGAGAGGCGCAGAGGGTCAATGCTGACTTTACCGGCGAT 540
Db      481 AACGTGACTGACGCGATCCTGAGAGGCGCAGAGGGTCAATGCTGACTTTACCGGCGAT 540
Qy      541 TATCAAAAGGGGTTTCCGGAACCTGGAACGGGTGCTTAATTTCCGCAATCAACTTGGC 600
Db      541 TATCAAAAGGGGTTTCCGGAACCTGGAACGGGTGCTTAATTTCCGCAATCAACTTGGC 600
Qy      601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGAGGCAATTCATCGGAATC 660
Db      601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGAGGCAATTCATCGGAATC 660
Qy      661 AAGGTGAGCGCGCATGTGTCTCATTTAACCGGTGCGGTAGCCTGCAATCAATGCTGACG 720
Db      661 AAGGTGAGCGCGCATGTGTCTCATTTAACCGGTGCGGTAGCCTGCAATCAATGCTGACG 720
Qy      721 GAGATATTTCTCTGCAACAGCACAAGGAATGCCGAGCGCGGTGTGAGATTGATCAAGCAGCG 780
Db      721 GAGATATTTCTCTGCAACAGCACAAGGAATGCCGAGCGCGGTGTGAGATTGATCAAGCAGCG 780
Qy      781 GATTCAACAGAGTGAACACCTGTGTAAGTTGCAATACGCGCAATTTGATTGCTACAA 840
Db      781 GATTCAACAGAGTGAACACCTGTGTAAGTTGCAATACGCGCAATTTGATTGCTACAA 840
Qy      841 CGCAGCGCAGAGGTTGCGCGGCGCGCAGCGCAGCGCGGTATTAAGATTGATCAAGCAGCG 900
Db      841 CGCAGCGCAGAGGTTGCGCGGCGCGCAGCGCAGCGCGGTATTAAGATTGATCAAGCAGCG 900
Qy      901 TTGAGCGCCCATCCACGCGGAAACAGCGGTATGAGTGAACATTCACCTTCAGTGTG 960
Db      901 TTGAGCGCCCATCCACGCGGAAACAGCGGTATGAGTGAACATTCACCTTCAGTGTG 960
Qy      961 TTTATGCGCGGACAGATTAATCTGCAATCTCGCGCGCGGCGCAGCTGAGCTCAACTCG 1020
Db      961 TTTATGCGCGGACAGATTAATCTGCAATCTCGCGCGCGGCGCAGCTGAGCTCAACTCG 1020
Qy      1021 ACGCTTCCCGTCAAGCCGATTAACAGCGCGCAGGTGATGATGCTGTTGAACGCTGCG 1080
Db      1021 ACGCTTCCCGTCAAGCCGATTAACAGCGCGCAGGTGATGATGCTGTTGAACGCTGCG 1080
Qy      1081 CGTGGGCTTAAGCGATTAACAGCGCAGGTGATGATGCTGCTGCTTCCAGACTTTACAG 1140
Db      1081 CGTGGGCTTAAGCGATTAACAGCGCAGGTGATGATGCTGCTGCTTCCAGACTTTACAG 1140

```

```

Qy      1141 CAGATGCGTATTAACGCGCGCTGTCATTAATTAACGCGCGCGAGAGTGAACCTGACC 1200
Db      1141 CAGATGCGTATTAACGCGCGCTGTCATTAATTAACGCGCGCGAGAGTGAACCTGACC 1200
Qy      1201 CTGGCAGATGTGAAGAGCGAATAATGCGCAGCGCATGTGTTGTTGGCAGGTTTACGCA 1260
Db      1201 CTGGCAGATGTGAAGAGCGAATAATGCGCAGCGCATGTGTTGTTGGCAGGTTTACGCA 1260
Qy      1261 ATCGTAATGAAGCAGCATACCGGCGGTGAGTTTGATCTCATC 1306
Db      1261 ATCGTAATGAAGCAGCATACCGGCGGTGAGTTTGATCTCATC 1306

RESULT 15
AD050303
ID      AD050303 standard; DNA, 1323 BP.
XX
AC      AD050303;
XX
DT      29-JUL-2004 (first entry)
XX
DE      Escherichia coli W phytase 875PH2 mutant DNA.
XX
KW      phytase; bacteria; thermal tolerance; protease stability; foodstuffs;
KW      phytase; animal feed; fish feed; dough; baking; mutant; gene; ds.
XX
OS      Escherichia coli.
XX
FH      Key
FT      CDS
FT      1..1323
FT      /tag= a
FT      /product= "Mutant phytase enzyme"
FT      214..216
FT      /tag= b
FT      /note= "Encodes Arg"
FT      mutation
FT      /tag= c
FT      replace(527,A)
FT      /tag= d
FT      replace(893,T)
FT      /tag= e
FT      mutation
FT      /tag= f
FT      replace(895,G)
XX
PN      US2004091968-A1.
XX
PD      13-MAY-2004.
XX
PF      20-JUN-2003; 2003US-00601319.
XX
PR      13-AUG-1997; 97US-00910798.
PR      01-MAR-1999; 99US-00259214.
PR      13-APR-1999; 99US-00291914.
PR      25-MAY-1999; 99US-00318528.
PR      25-MAY-2000; 2000US-00580515.
PR      24-MAY-2001; 2001US-00866379.
XX
PA      (SHOR/) SHORT J M.
PA      (KRETZ) KRETZ K.
PA      (GRAY/) GRAY K A.
PA      (BART/) BARTON N R.
PA      (GARR/) GARRETT J B.
PA      (ODON/) O'DONOGHUE E.
PA      (MATH/) MATHER E J.
XX
PI      Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E,
PI      Mather EJ;
XX
XX      WPI; 2004-374952/35.
XX      DR      P-PSDB; AD050292; AD050304.
XX      XX
XX      Producing phytase, involves providing nucleic acid derived from bacteria

```

PT encoding polypeptide having phytase activity, and expressing nucleic acid
PT in yeast.

PS Disclosure; Page, 74pp; English.

XX
CC The invention relates to a method for producing phytase that involves
CC providing a nucleic acid encoding phytase derived from a bacterium, and
CC expressing the nucleic acid in a yeast under conditions that allow
CC expression of the enzyme in the yeast. The invention also relates to
CC modified phytase enzyme which has improved thermal tolerance and protease
CC stability at low pH. The phytase enzyme can be used in foodstuffs to
CC improve the feeding value of phytate rich ingredients, and in diet of
CC numerous animals including mammals, fowls and fishes, commercially
CC significant mammals such as pigs, goats, laboratory rodents, commercially
CC significant avian species such as chicken, ducks, doves, parrot, etc.,
CC commercially farmed fish such as guppy, zebrafish, mollus, swordtail,
CC etc., in dough making and baking, in dietary aids for animals. The method
CC provides easy manufacture of the active ingredient loaded biocompatible
CC composition, higher yields and loading efficiency. The phytase
CC incorporated in to the dietary aid is safe for animals. The present
CC sequence is Escherichia coli W phytase mutant DNA. Note: The present
CC sequence is not shown in the specification but has been derived from
CC Escherichia coli B phytase DNA AD050291.

XX
SQ Sequence 1323 BP; 324 A; 353 C; 356 G; 289 T; 0 U; 1 Other;

Query Match 96.7%; Score 1265; DB 12; Length 1323;

Best Local Similarity 98.0%; Pred. No. 0;

Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 ATGAAGGAGATCTTATCCATTTTATCTCTTGAATCCGTTAAACCCCGAATCTGA 60
DB 1 ATGAAGGAGATCTTATCCATTTTATCTCTTGAATCCGTTAAACCCCGAATCTGA 60
QY 61 TTGCTCAGAGTGAAGCCGAGCTGAAGCTGAAGAGTGTGATTTGATGCTGATGAT 120
DB 61 TTGCTCAGAGTGAAGCCGAGCTGAAGCTGAAGAGTGTGATTTGATGCTGATGAT 120
QY 121 GTGCTGCTCCAAAGCCGCAAGCACTGATGCAAGATGTCACCCCGAAGCATGCGCA 180
DB 121 GTGCTGCTCCAAAGCCGCAAGCACTGATGCAAGATGTCACCCCGAAGCATGCGCA 180
QY 181 ACCTGGCCGGTAAACCTGGGTGAGCTGACACCGCCGGTGTGAGCTAATGCCATCTGC 240
DB 181 ACCTGGCCGGTAAACCTGGGTGAGCTGACACCGCCGGTGTGAGCTAATGCCATCTGC 240
QY 241 GGACATTAATGCGCTGAGCGTCTGCTGAGCCGAGATTGCTCTTAAATGTGGTCCCG 300
DB 241 GGACATTAATGCGCGCAAGGCTGCTGCTGAGCCGAGATTGCTCTTAAATGTGGTCCCG 300
QY 301 CAGTCTGGTCAAGGTGCGGATTATTTGCTGATGTCAGACGAGCTTACCGTTAAACAGGCGAA 360
DB 301 CAGTCTGGTCAAGGTGCGGATTATTTGCTGATGTCAGACGAGCTTACCGTTAAACAGGCGAA 360
QY 361 GCTTGGCCGGCCGGCTGAGCACTGATGCAATTAACCTGATCAATCCAGGCAATAG 420
DB 361 GCTTGGCCGGCCGGCTGAGCACTGATGCAATTAACCTGATCAATCCAGGCAATAG 420
QY 421 TCCAGTCCCGATCCGTTATTTATCTCTTAAACCTGGGCTTTGCAATGGAATACGCG 480
DB 421 TCCAGTCCCGATCCGTTATTTATCTCTTAAACCTGGGCTTTGCAATGGAATACGCG 480
QY 481 AACGTACTGACGCGATCTCTGAGAGGCGAGAGGGTCAATTGCTGACTTTAACCGGGCAT 540
DB 481 AACGTACTGACGCGATCTCTGAGAGGCGAGAGGGTCAATTGCTGACTTTAACCGGGCAT 540
QY 541 TATCAAGGCGGCTTGGGAACTGGAACGGGTCTTAATTTCCGCAATCAAACTTGTGC 600
DB 541 TATCAAGGCGGCTTGGGAACTGGAACGGGTCTTAATTTCCGCAATCAAACTTGTGC 600
QY 601 CTTAAACGTGAGAAACAGGACGAAGCTGTTCAATTAACGACGAGCATTAACATCGGAATC 660
DB 601 CTTAAACGTGAGAAACAGGACGAAGCTGTTCAATTAACGACGAGCATTAACATCGGAATC 660

QY 661 AAGTGAAGCCCGGATGCTGTCTCATTTAACCGGTGCGTTAAGCTTGCATCAATGCTGACG 720
DB 661 AAGTGAAGCCCGGATGCTGTCTCATTTAACCGGTGCGTTAAGCTTGCATCAATGCTGACG 720
QY 721 GAGATATTTCTCTGCAACAGCAAGGAAATGCGGAGCCGGGGTGGGGAAGATCAACC 780
DB 721 GAGATATTTCTCTGCAACAGCAAGGAAATGCGGAGCCGGGGTGGGGAAGATCAACC 780
QY 781 GATTACACACAGTGAACACCTTGTGAAGTTGATTAACGCGCAATTTGATTGCTACAA 840
DB 781 GATTACACACAGTGAACACCTTGTGAAGTTGATTAACGCGCAATTTGATTGCTACAA 840
QY 841 CGCAGCCGAGAGTTGCCCGGACCGCCGCAACCCCGTTATTAATGATTAATCAAGACAGCG 900
DB 841 CGCAGCCGAGAGTTGCCCGGACCGCCGCAACCCCGTTATTAATGATTAATCAAGACAGCG 900
QY 901 TTGAAGCCCGCATCCACCGGCAAAAACAGGCGTATGTTGACATTAACCACTTCACTGCTG 960
DB 901 TTGAAGCCCGCATCCACCGGCAAAAACAGGCGTATGTTGACATTAACCACTTCACTGCTG 960
QY 961 TTTATGCGCGGACACAGTACTAATCTGCAATCTCGGCGGCGCACTGGAAGCTCAACTGG 1020
DB 961 TTTATGCGCGGACACAGTACTAATCTGCAATCTCGGCGGCGCACTGGAAGCTCAACTGG 1020
QY 1021 ACGTTCCCGGTCAAGCCGATTAACACGCGCCAGGTGTTGAATGTTGAAACGCTGG 1080
DB 1021 ACGTTCCCGGTCAAGCCGATTAACACGCGCCAGGTGTTGAATGTTGAAACGCTGG 1080
QY 1081 CGTGGCTTAAGCATTAACAGCAGTGAATTCAGTTTGTGCTGCTTTCAGACTTTACAG 1140
DB 1081 CGTGGCTTAAGCATTAACAGCAGTGAATTCAGTTTGTGCTGCTTTCAGACTTTACAG 1140
QY 1141 CAGATGCGTGAATTAACCGCCTGTCATTAATACGCGCCGAGAGATGTAACCTGACC 1200
DB 1141 CAGATGCGTGAATTAACCGCCTGTCATTAATACGCGCCGAGAGATGTAACCTGACC 1200
QY 1201 CTGGCAGATGTAAGACGCAAAATGCGAGGCGCATGTGTGCTGGCAGGTTTACGCA 1260
DB 1201 CTGGCAGATGTAAGACGCAAAATGCGAGGCGCATGTGTGCTGGCAGGTTTACGCA 1260
QY 1261 ATCGTAATGAAGCAGCATACCGCGCTGCAAGTTTGAATCTCAATC 1306
DB 1261 ATCGTAATGAAGCAGCATACCGCGCTGCAAGTTTGAATCTCAATC 1306

Search completed: June 14, 2006, 10:31:29
Job time : 893.897 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2006, 10:00:55 ; Search time 7316.05 Seconds

(without alignments) 11432.839 Million cell updates/sec

Title: US-10-601-319-9

Perfect score: 1308

Sequence: 1 atgaagcgacatcatcc.....gcatttgagatcatcata 1308

Scoring table: IDENTITY_NUC

Searched: 636136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmb1.*
1: gb_env:*
2: gb_pat:*
3: gb_ph:*
4: gb_pl:*
5: gb_pr:*
6: gb_ro:*
7: gb_ses:*
8: gb_sy:*
9: gb_un:*
10: gb_vl:*
11: gb_ov:*
12: gb_hrg:*
13: gb_in:*
14: gb_cm:*
15: gb_da:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1265	96.7	1323	2	AR108133 Sequence
2	1265	96.7	1323	2	AR127818 Sequence
3	1265	96.7	1323	2	AR130956 Sequence
4	1265	96.7	1323	2	AR492966 Sequence
5	1265	96.7	1323	2	AR636185 Sequence
6	1265	96.7	1323	2	AX052713 Sequence
7	1265	96.7	1323	2	AX338966 Sequence
8	1265	96.7	1323	2	AX356566 Sequence
9	1264.8	96.7	110000	15	U00096.10 Continuation (11 o
10	1264.8	96.7	110000	15	AP009048_10 Continuation (11 o
11	1263.8	96.6	1901	2	AR636189 Sequence
12	1263.8	96.6	1901	2	AR636190 Sequence
13	1263.8	96.6	1901	2	AX356571 Sequence
14	1263.8	96.6	1901	2	AX356572 Sequence
15	1263.8	96.6	1901	15	ECOACPPAA MS8708 Becherichia
16	1263.2	96.5	110000	15	CP000038_10 Continuation (11 o
17	1263.2	96.5	1901	2	AR636188 Sequence
18	1263.2	96.5	1901	2	AR636191 Sequence

19	1262.2	96.5	1901	2	AX356570 Sequence
20	1262.2	96.5	1901	2	AX356574 Sequence
21	1262.2	96.5	1901	15	ECOACPPHA L03371 Becherichia
22	1261.6	96.5	1315	15	AF537219 Becherichia
23	1260.6	96.4	1901	15	ECOACPHC L03372 Becherichia
24	1260.6	96.4	1901	15	ECOACPHD L03373 Becherichia
25	1260.6	96.4	1901	15	ECOACPHF L03375 Becherichia
26	1260.6	96.4	1901	15	ECOACPHI L03376 Becherichia
27	1259.2	96.3	3470	2	AX042376 Sequence
28	1259.2	96.3	4060	2	AX042374 Sequence
29	1259.2	96.3	6116	2	AX042375 Sequence
30	1259.2	96.3	6708	2	AX042373 Sequence
31	1259.2	96.3	17732	2	AX042378 Sequence
32	1259.2	96.3	20623	2	AX042372 Sequence
33	1259	96.3	1901	15	ECOACPHA L03370 Becherichia
34	1246.2	95.2	5421	2	AX042377 Sequence
35	1245.6	95.2	110000	15	CP000036_22 Continuation (23 o
36	1239.2	94.7	110000	15	AB005674_10 Continuation (11 o
37	1237.6	94.6	110000	15	AB014073_10 Continuation (11 o
38	1233.6	94.3	1486	2	AR630587 Sequence
39	1228	93.9	1489	2	AR277753 Sequence
40	1228	93.9	1489	15	AY496073 Becherichia
41	1222.4	93.5	1486	2	AR630588 Sequence
42	1218.8	93.2	110000	15	AB005174_12 Continuation (13 o
43	1218.8	93.2	110000	15	AB005174_13 Continuation (13 o
44	1218.8	93.2	110000	15	BA000007_12 Continuation (9 of
45	1218.8	93.2	110000	15	CP000034_08 Continuation (9 of

ALIGNMENTS

RESULT 1	AR108133	1323 bp	DNA	linear	PAT 14-FEB-2001
LOCUS	AR108133	Sequence 1 from patent US 6110719.			
DEFINITION	AR108133				
ACCESSION	AR108133				
VERSION	AR108133.1	GI:12823620			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 1323)				
AUTHORS	Kretz, K.				
TITLE	Phytase				
JOURNAL	Patent: US 6110719-A 1 29-AUG-2000;				
FEATURES	Location/Qualifiers				
source	1..1323				
	/organism="unknown"				
	/mol_type="unassigned DNA"				

ORIGIN	Query Match	96.7%; Score 1265; DB 2; Length 1323;
	Best Local Similarity	98.0%; Pred. No. 0;
	Matches 1280; Conservative	0; Mismatches 26; Indels 0; Gaps 0;
QY	1	ATGAAAGGATCTTAATCCATTTTATCTTCTGATTCGGTTAACCCGCAATCTGCA 60
DB	1	ATGAAAGGATCTTAATCCATTTTATCTTCTGATTCGGTTAACCCGCAATCTGCA 60
QY	61	TTTCCTCAAGTGAAGCCGAGACTGAAGTGAAGTGGTGGATTTGATTCGTCATAGT 120
DB	61	TTTCCTCAAGTGAAGCCGAGACTGAAGTGAAGTGGTGGATTTGATTCGTCATAGT 120
QY	121	GTGGCTGCTCCAAACCAAGCCGACCACTGATGACAGATGTCACCCGACGATGGCCA 180
DB	121	GTGGCTGCTCCAAACCAAGCCGACCACTGATGACAGATGTCACCCGACGATGGCCA 180
QY	181	ACCTGGCCGGTAAACTGGTGAAGTGAACACCGCGCGGTGGTGAAGTAACTGCTATCTC 240
DB	181	ACCTGGCCGGTAAACTGGTGAAGTGAACACCGCGCGGTGGTGAAGTAACTGCTATCTC 240
QY	241	GGACATTACTGGCGCTCAGCGCTCTGGTGAAGCAGGATTCGCTTAATGTGGTCCCGC 300

```
|||||
Db 241 GGAATTACCAACGCGACGCTCTGCTAGCCGACGGAATTGCTGGGAAAAAGGCGTCCCG 300
Qy 301 CAGTCTGTCAGGTTCGCGATTATGCTGATGTCCAGACGCTACCCGTAACACGGCGAA 360
Db 301 CAGTCTGTCAGGTTCGCGATTATGCTGATGTCCAGACGCTACCCGTAACACGGCGAA 360
Qy 361 GCCTTCGCGCGCGGCTGGGACCTGACTGTGCAATACCGTACATACCAGGAGATACG 420
Db 361 GCCTTCGCGCGCGGCTGGGACCTGACTGTGCAATACCGTACATACCAGGAGATACG 420
Qy 421 TCCAGTCCCGATCCGTTATTTAATCTCTTAAAAACTGGCGTTTGCCAATGATACGCG 480
Db 421 TCCAGTCCCGATCCGTTATTTAATCTCTTAAAAACTGGCGTTTGCCAATGATACGCG 480
Qy 481 AACGTGACTGACGGATCTCTGAGAGGGCAGAGGGGTCAATTGCTGACTTTACGGGGCAT 540
Db 481 AACGTGACTGACGGATCTCTGAGAGGGCAGAGGGGTCAATTGCTGACTTTACGGGGCAT 540
Qy 541 TATCAACGCGGCTTTCGGAACCTGGAACGGGTGCTTAATTTCCGCAATCAACTGTGC 600
Db 541 CCGCAACGCGGCTTTCGGAACCTGGAACGGGTGCTTAATTTCCGCAATCAACTGTGC 600
Qy 601 CTTAAACGTGAGAAACAGACGAAAGCTGTTCAATTACGACGATTAACATCCGAACTC 660
Db 601 CTTAAACGTGAGAAACAGACGAAAGCTGTTCAATTACGACGATTAACATCCGAACTC 660
Qy 661 AAGGTAGCGCCGACGTGTCTCATTTAACCGGTGCGTAAAGCTTCGACATGCTGACG 720
Db 661 AAGGTAGCGCCGACGTGTCTCATTTAACCGGTGCGTAAAGCTTCGACATGCTGACG 720
Qy 721 GAGATATTTCTCTGCAACAGACACAGGGAATGCGGAGCGGGGTGGGAAAGATCAC 780
Db 721 GAGATATTTCTCTGCAACAGACACAGGGAATGCGGAGCGGGGTGGGAAAGATCAC 780
Qy 781 GATTTCACACAGTGAACACCTTGTAAAGTTGCAATACGCGCAATTTGATTTGCTACAA 840
Db 781 GATTTCACACAGTGAACACCTTGTAAAGTTGCAATACGCGCAATTTTATTTGCTACAA 840
Qy 841 CGCACGCGCAGAGGTGCGCCGACGCGCCACCCCGTATTAAGATTGATCAACACGCG 900
Db 841 CGCACGCGCAGAGGTGCGCCGACGCGCCACCCCGTATTAAGATTGATCAACACGCG 900
Qy 901 TTGACGCGCCATCCACCGCAAAAACAGGCGTATGATGACATTAACCACTTCAGTACTG 960
Db 901 TTGACGCGCCATCCACCGCAAAAACAGGCGTATGATGACATTAACCACTTCAGTACTG 960
Qy 961 TTATTCGCGGACAGATTAATCTGCAAAATCTCGGCGCGCACTGGAGCTCAACTCG 1020
Db 961 TTATTCGCGGACAGATTAATCTGCAAAATCTCGGCGCGCACTGGAGCTCAACTCG 1020
Qy 1021 ACGCTTCCCGGTCAAGCCGGAATACACGCGCAGGTGATGAATCTGTGTGGAACGCTCG 1080
Db 1021 ACGCTTCCCGGTCAAGCCGGAATACACGCGCAGGTGATGAATCTGTGTGGAACGCTCG 1080
Qy 1081 CGTGGGCTAACGATTAACGCAAGTGAATTCAGGTTTCGCTGTCTTCCAACTTTACAG 1140
Db 1081 CGTGGGCTAACGATTAACGCAAGTGAATTCAGGTTTCGCTGTCTTCCAACTTTACAG 1140
Qy 1141 CAGATGCGTGAATAAACGCGCTGTCAATTAATAACGCGCCGAGAGAGTGAACCTGACC 1200
Db 1141 CAGATGCGTGAATAAACGCGCTGTCAATTAATAACGCGCCGAGAGAGTGAACCTGACC 1200
Qy 1201 CTGGCAGATGTAAGAGACGAATGCGACGGGCATGTGTGTTGCGAGGTTTAAACGAA 1260
Db 1201 CTGGCAGATGTAAGAGACGAATGCGACGGGCATGTGTGTTGCGAGGTTTAAACGAA 1260
Qy 1261 ATCGTAATGAAGACGATACCGGCGTGCAGTTTGAGATCTCATC 1306
Db 1261 ATCGTAATGAAGACGATACCGGCGTGCAGTTTGAGATCTCATC 1306
```

RESULT 2

```
AR127818
LOCUS AR127818 1323 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6183740.
ACCESSION AR127818
VERSION AR127818.1 GI:14115480
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1323)
AUTHORS Short,J.M. and Kretz,K.A.
TITLE Recombinant bacterial phylases and uses thereof
JOURNAL Patent: US 6183740-A 1 06-FEB-2001;
FEATURES
location/Qualifiers
1..1323
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 96.7%; Score 1265; DB 2; Length 1323;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
Qy 1 ATGAAACGATCTTAATCCCATTTTATCTTCTGATTCGGTTAACCCGCAATCTGCA 60
Db 1 ATGAAACGATCTTAATCCCATTTTATCTTCTGATTCGGTTAACCCGCAATCTGCA 60
Qy 61 TTGCTCAGAGTGAACCGGACCTGAGAGCTGGAAGTGTGATTTGCACTGTCATGCT 120
Db 61 TTGCTCAGAGTGAACCGGACCTGAGAGCTGGAAGTGTGATTTGCACTGTCATGCT 120
Qy 121 GTGCGTCTCAACCAAGGCGACGCAACTGATGAGGATGTCAACCCGACGATGGCCA 180
Db 121 GTGCGTCTCAACCAAGGCGACGCAACTGATGAGGATGTCAACCCGACGATGGCCA 180
Qy 181 ACCTGGCGGTAAACTGGGTGAGCTGACACGCGCGGTGTGAGCTTAATCGCTATCTC 240
Db 181 ACCTGGCGGTAAACTGGGTGAGCTGACACGCGCGGTGTGAGCTTAATCGCTATCTC 240
Qy 241 GGCATTAATCTGGGTGAGCGCTGTGTAAGCCGAGAGTGTGCTTAATGTGCTGCGCG 300
Db 241 GGCATTAATCTGGGTGAGCGCTGTGTAAGCCGAGAGTGTGCTTAATGTGCTGCGCG 300
Qy 301 CAGTCTGTCAGGTTCGCGATTATGCTGATGTCCAGACGCTACCCGTAACACGGCGAA 360
Db 301 CAGTCTGTCAGGTTCGCGATTATGCTGATGTCCAGACGCTACCCGTAACACGGCGAA 360
Qy 361 GCCTTCGCGCGCGGCTGGGACCTGACTGTGCAATTAACGTTACCAACGAGATACG 420
Db 361 GCCTTCGCGCGCGGCTGGGACCTGACTGTGCAATTAACGTTACCAACGAGATACG 420
Qy 421 TCCAGTCCCGATCCGTTATTTAATCTCTTAAAAACTGGCGTTTGCCAATGATACGCG 480
Db 421 TCCAGTCCCGATCCGTTATTTAATCTCTTAAAAACTGGCGTTTGCCAATGATACGCG 480
Qy 481 AACGTGACTGACGGATCTCTGAGAGGGCAGAGGGTCAATTGCTGACTTTACGGGCAT 540
Db 481 AACGTGACTGACGGATCTCTGAGAGGGCAGAGGGTCAATTGCTGACTTTACGGGCAT 540
Qy 541 TATCAACGCGGCTTTCGGAACCTGGAACGGGTGCTTAATTTTCGCAATCAACTGTGC 600
Db 541 CCGCAACGCGGCTTTCGGAACCTGGAACGGGTGCTTAATTTTCGCAATCAACTGTGC 600
Qy 601 CTTAAACGTGAGAAACAGACGAAAGCTGTTCAATTACGACGATTAACATCCGAACTC 660
Db 601 CTTAAACGTGAGAAACAGACGAAAGCTGTTCAATTACGACGATTAACATCCGAACTC 660
Qy 661 AAGGTAGCGCCGACGTGTCTCATTTAACCGGTGCGTAAAGCTTCGACATGCTGACG 720
Db 661 AAGGTAGCGCCGACATGTCTCATTTAACCGGTGCGTAAAGCTTCGACATGCTGACG 720
Qy 721 GAGATATTTCTCTGCAACAGACACAGGAATGCGGAGCTGGGGTGGGAAAGATCAC 780
Db 721 GAGATATTTCTCTGCAACAGACACAGGAATGCGGAGCTGGGGTGGGAAAGATCAC 780
```

Db 721 GAGATATTTCTCTGCAACAGCAAGGAATGCGAGCCGGGGTGGGAAAGATCAC 780
Qy 781 GATTGACACAGATGGAACACTTGTCTAATTGTCATTAAGCGGAATTTGATTGCTACAA 840
Db 781 GATTGACACAGATGGAACACTTGTCTAATTGTCATTAAGCGGAATTTGATTGCTACAA 840
Qy 841 CGCAGCCGAGAGTTGCGCGAGCGCGCAACCCGTTATTAGATTGATCAAGACAGG 900
Db 841 CGCAGCCGAGAGTTGCGCGAGCGCGCAACCCGTTATTAGATTGATCAAGACAGG 900
Qy 901 TTGAGCGCCCATCCACCGCAAAAACAGCGGTATGGTGACATTAACCACTTCAGTCTG 960
Db 901 TTGAGCGCCCATCCACCGCAAAAACAGCGGTATGGTGACATTAACCACTTCAGTCTG 960
Qy 961 TTTATGCGCGGACACGATCTAATCTGCGAATCTTGCGCGCGCACTGAGCTCAACTG 1020
Db 961 TTTATGCGCGGACACGATCTAATCTGCGAATCTTGCGCGCGCACTGAGCTCAACTG 1020
Qy 1021 ACGCTTCCCGGTCAAGCGCGATTAACAGCGCGCGAGGTGTAACCTGGTGTGAAAGCTGG 1080
Db 1021 ACGCTTCCCGGTCAAGCGCGATTAACAGCGCGCGAGGTGTAACCTGGTGTGAAAGCTGG 1080
Qy 1081 CGTGGCTTAAGCGATTAACAGCGAGTGAATTCAGTTTCGCTGCTTCCAGACTTTACAG 1140
Db 1081 CGTGGCTTAAGCGATTAACAGCGAGTGAATTCAGTTTCGCTGCTTCCAGACTTTACAG 1140
Qy 1141 CAGATGCGGTGAATAACCGCGCTGTCTAATAATACCGCGCGGAGAGGTGAACCTGACC 1200
Db 1141 CAGATGCGGTGAATAACCGCGCTGTCTAATAATAACCGCGCGGAGAGGTGAACCTGACC 1200
Qy 1201 CTGGCAGATGTGAAGAGCGGAATGCGAGGCAATGTTTCTGGCAGATTTCAGCAA 1260
Db 1201 CTGGCAGATGTGAAGAGCGGAATGCGAGGCAATGTTTCTGGCAGATTTCAGCAA 1260
Qy 1261 ATCGTAATGAAGACGCGATACCGCGCGTGCAGTTTGAGATCTCATC 1306
Db 1261 ATCGTAATGAAGACGCGATACCGCGCGTGCAGTTTGAGATCTCATC 1306

RESULT 3
ARI30956 1323 bp DNA linear PAT 16-MAY-2001
LOCUS AR130956
DEFINITION Sequence 1 from patent US 6190897.
ACCESSION AR130956
VERSION AR130956.1 GI:14119281
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1323)
AUTHORS Kretz, K.
TITLE Phytase
JOURNAL Patent: US 6190897-A 1 20-FEB-2001;
FEATURES Location/Qualifiers
source 1..1323
/organism="Unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 96.7%; Score 1265; DB 2; Length 1323;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 1 ATGAAAGGATCTTAATCCATTTTATCTCTTGATTCGCTTAACCCCGCAATCTGCA 60
Db 1 ATGAAAGGATCTTAATCCATTTTATCTCTTGATTCGCTTAACCCCGCAATCTGCA 60
Qy 61 TTGCGTCAAGTGAAGCGGAGCTGAAGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 120
Db 61 TTGCGTCAAGTGAAGCGGAGCTGAAGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 120
Qy 121 GTGCGTGTCTCAACCAAGGCCACGCACTGATGACAGATGTCAACCCAGACGATGCGCA 180
Db 121 GTGCGTGTCTCAACCAAGGCCACGCACTGATGACAGATGTCAACCCAGACGATGCGCA 180

Db 121 GTGCGTGTCTCAACCAAGGCCACGCACTGATGACAGATGTCAACCCAGACGATGCGCA 180
Qy 181 ACCTGCCGGTAAACTGGGTGAGCTGACACCGCGCGGTGGTGAAGTAAATGCTTATCTC 240
Db 181 ACCTGCCGGTAAACTGGGTGAGCTGACACCGCGCGGTGGTGAAGTAAATGCTTATCTC 240
Qy 241 GGAATTAATGCGCGTCAAGGTCTGATGACCGGATTTGCTGCGCTTAATGTTGGCTGCGCG 300
Db 241 GGAATTAATGCGCGTCAAGGTCTGATGACCGGATTTGCTGCGCTTAATGTTGGCTGCGCG 300
Qy 301 CAGTCTGTGACAGTGCAGATTATTTGCTGATGTCAGACAGCGTAAACCGTAAACAGCGCA 360
Db 301 CAGTCTGTGACAGTGCAGATTATTTGCTGATGTCAGACAGCGTAAACCGTAAACAGCGCA 360
Qy 361 GCTTTCGCGCGCGGCTGCGCACTGATGTCGATTAACCGTAAACCGTAAACCGGACAGTACG 420
Db 361 GCTTTCGCGCGCGGCTGCGCACTGATGTCGATTAACCGTAAACCGTAAACCGGACAGTACG 420
Qy 421 TCCAGTCCCGAATCCGTTAATTTAATCTCTTAATAAATGCGCGTTTGGCACTGATTAACCG 480
Db 421 TCCAGTCCCGAATCCGTTAATTTAATCTCTTAATAAATGCGCGTTTGGCACTGATTAACCG 480
Qy 481 AACGTGACTGAACGCGATCTCGAGAGGCGAGAGGCTCAATTGCTGACTTTACCGGCGAT 540
Db 481 AACGTGACTGAACGCGATCTCGAGAGGCGAGAGGCTCAATTGCTGACTTTACCGGCGAT 540
Qy 541 TATCAACCGCGCTTTCGGAACCTGGAACCGGCTGTTAATTTTCGCAATCAAACTTTGTC 600
Db 541 CGGCAAAAGCGGCTTTCGGAACCTGGAACCGGCTGTTAATTTTCGCAATCAAACTTTGTC 600
Qy 601 CTTAAACGTGAAGAAACAGAGAGAAAGCTTCAATTAAAGCGAGCACTTAAACATCGGAATC 660
Db 601 CTTAAACGTGAAGAAACAGAGAGAAAGCTTCAATTAAAGCGAGCACTTAAACATCGGAATC 660
Qy 661 AAGGTGACCGCGCACTGTGTCTAATTAAACGGTGGCTGAAGCTGCGCATTAATGCTGACG 720
Db 661 AAGGTGACCGCGCACTGTGTCTAATTAAACGGTGGCTGAAGCTGCGCATTAATGCTGACG 720
Qy 721 GAGATATTTCTCTGCAACAGCAAGGAATGCGGAGCCGGGGTGGGAAAGATCAC 780
Db 721 GAGATATTTCTCTGCAACAGCAAGGAATGCGGAGCCGGGGTGGGAAAGATCAC 780
Qy 781 GATTGACACAGATGGAACACTTGTCTAATTGTCATTAAGCGCAATTTGATTGCTACAA 840
Db 781 GATTGACACAGATGGAACACTTGTCTAATTGTCATTAAGCGCAATTTGATTGCTACAA 840
Qy 841 CGCAGCCGAGAGTTGCGCGAGCGCGCAACCCGTTATTAGATTGATCAAGACAGG 900
Db 841 CGCAGCCGAGAGTTGCGCGAGCGCGCAACCCGTTATTAGATTGATCAAGACAGG 900
Qy 901 TTGAGCGCCCATCCACCGCAAAAACAGCGGTATGGTGACATTAACCACTTCAAGTACG 960
Db 901 TTGAGCGCCCATCCACCGCAAAAACAGCGGTATGGTGACATTAACCACTTCAAGTACG 960
Qy 961 TTTATGCGCGGACACGATCTAATCTGCGAATCTTGCGCGCGCACTGAGCTCAACTG 1020
Db 961 TTTATGCGCGGACACGATCTAATCTGCGAATCTTGCGCGCGCACTGAGCTCAACTG 1020
Qy 1021 ACGCTTCCCGGTCAAGCGCGATTAACAGCGCGCGAGGTGTAACCTGGTGTGAAAGCTGG 1080
Db 1021 ACGCTTCCCGGTCAAGCGCGATTAACAGCGCGCGAGGTGTAACCTGGTGTGAAAGCTGG 1080
Qy 1081 CGTGGCTTAAGCGATTAACAGCGAGTGAATTCAGTTTCGCTGCTTCCAGACTTTACAG 1140
Db 1081 CGTGGCTTAAGCGATTAACAGCGAGTGAATTCAGTTTCGCTGCTTCCAGACTTTACAG 1140
Qy 1141 CAGATGCGGTGAATAACCGCGCTGTCTAATAATAACCGCGCGGAGAGGTGAACCTGACC 1200
Db 1141 CAGATGCGGTGAATAACCGCGCTGTCTAATAATAACCGCGCGGAGAGGTGAACCTGACC 1200
Qy 1201 CTGGCAGATGTGAAGAGCGGAATGCGAGGCAATGTTTCTGGCAGATTTCAGCAA 1260
Db 1201 CTGGCAGATGTGAAGAGCGGAATGCGAGGCAATGTTTCTGGCAGATTTCAGCAA 1260

QY 1261 ATCGTGAATGAAGCAGCATACCGGCGGTGAGTTGAGATCTCATC 1306
Db 1261 ATCGTGAATGAAGCAGCATACCGGCGGTGAGTTGAGATCTCATC 1306

RESULT 4
AR492966
LOCUS AR492966 1323 bp DNA linear PAT 15-MAY-2004
DEFINITION Sequence 1 from patent US 6720014.
ACCESSION AR492966
VERSION AR492966.1 GI:47264255
KEYWORDS
SOURCE .
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1323)
AUTHORS Short,J.M. and Kretz,K.A.
TITLE Phytase-containing foodstuffs and methods of making and using them
JOURNAL Patent: US 6720014-A 1 13-APR-2004;
Diversa Corporation; San Diego, CA

FEATURES
source 1..1323
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 96.7%; Score 1265; DB 2; Length 1323;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 ATGAAAGCATCTTAATCCCATTTTATCTCTGATTCGCTTAACCCGCGCATCTGCA 60
Db 1 ATGAAAGCATCTTAATCCCATTTTATCTCTGATTCGCTTAACCCGCGCATCTGCA 60
QY 61 TTGCGCTCAGAGTGAAGCCGAGAGCTGAGAAAGTGTGATTCAGTCTCATGCT 120
Db 61 TTGCGCTCAGAGTGAAGCCGAGAGCTGAGAAAGTGTGATTCAGTCTCATGCT 120
QY 121 GTGCTGTCTCCACCAAGGCGCAAGCACTGATGAGATTCACCCGAGCGCATGGCCA 180
Db 121 GTGCTGTCTCCACCAAGGCGCAAGCACTGATGAGATTCACCCGAGCGCATGGCCA 180
QY 121 GTGCTGTCTCCACCAAGGCGCAAGCACTGATGAGATTCACCCGAGCGCATGGCCA 180
Db 121 GTGCTGTCTCCACCAAGGCGCAAGCACTGATGAGATTCACCCGAGCGCATGGCCA 180
QY 181 ACCGCGCGGTTAAACTGGGTGAGCTGACACCGCGCGGTGGTGAAGCTAATCGCTATCTC 240
Db 181 ACCGCGCGGTTAAACTGGGTGAGCTGACACCGCGCGGTGGTGAAGCTAATCGCTATCTC 240
QY 241 GAGCATTAATGCGCTGACCGCTGTGTAGCCGAGCGAATTTGCTTAATGTGCTGCCG 300
Db 241 GAGCATTAATGCGCTGACCGCTGTGTAGCCGAGCGAATTTGCTGTGCGAAGGCGCTGCCG 300
QY 301 CAGCTGTGTGATGAGTGTGATTCGATGATTCGACGAGCTACCGGTTAAACAGGCGGAA 360
Db 301 CAGCTGTGTGATGAGTGTGATTCGATGATTCGACGAGCTACCGGTTAAACAGGCGGAA 360
QY 361 GCGTTGCGCGCGCGGCTGAGCACTGATGTCGTAATACCGGATTCGCTTAATGTGCTGCCG 420
Db 361 GCGTTGCGCGCGCGGCTGAGCACTGATGTCGTAATACCGGATTCGCTTAATGTGCTGCCG 420
QY 421 TCCAGTCCCGATCCGTTATTTAATCTCTTAAACCTGCGCTTTGCCAATGAGATTAACGCG 480
Db 421 TCCAGTCCCGATCCGTTATTTAATCTCTTAAACCTGCGCTTTGCCAATGAGATTAACGCG 480
QY 481 AAGCTGATGACGAGATCTCTGAGAGGCGAGAGGCTCAATTTGCTGACTTTAACGGGCGAT 540
Db 481 AAGCTGATGACGAGATCTCTGAGAGGCGAGAGGCTCAATTTGCTGACTTTAACGGGCGAT 540
QY 541 TATCAACGAGGCTTTGCGGAACTGAAACGAGGCTGTAATTTTCGCAATCAAACTTGTGC 600
Db 541 CCGGAAACGAGGCTTTGCGGAACTGAAACGAGGCTGTAATTTTCGCAATCAAACTTGTGC 600
QY 601 CTTAAACGTGAGAAACAGAGCAAGGCTGTTCAATTAACGAGGATTAACCATCGAACTC 660
Db 601 CTTAAACGTGAGAAACAGAGCAAGGCTGTTCAATTAACGAGGATTAACCATCGAACTC 660

Db 601 CTTAAACGTGAGAAACAGAGCAAGGCTGTTCAATTAACGAGGATTAACCATCGAACTC 660
QY 661 AAGGTAGGCGCCGACCTGTGTCTTATTAACCGGTGGGTAGCCTTCGATCAATGCTGACG 720
Db 661 AAGGTAGGCGCCGACCAATGTCTTATTAACCGGTGGGTAGCCTTCGATCAATGCTGACG 720
QY 721 GAGATATTTCTCTGCAACAGCAGAGGATCCGAGCGGCGGTGGGAGGATCAACC 780
Db 721 GAGATATTTCTCTGCAACAGCAGAGGATCCGAGCGGCGGTGGGAGGATCAACC 780
QY 781 GATTCAACCAAGTGAACACCTTGTCTAAGTTTGATTAACCGGCAATTTGTTGCTACAA 840
Db 781 GATTCAACCAAGTGAACACCTTGTCTAAGTTTGATTAACCGGCAATTTGTTGCTACAA 840
QY 841 CGCAGCGCAAGGTTGGCCGCGCAGCCGCGCAACCCGCTTATTAATTAAGTGAACAGCGC 900
Db 841 CGCAGCGCAAGGTTGGCCGCGCAGCCGCGCAACCCGCTTATTAATTAAGTGAACAGCGC 900
QY 901 TTGACGCGCCATCAACCGCAAAAACAGGCTATGTGTGACATTAACCACTTCAGTACTG 960
Db 901 TTGACGCGCCATCAACCGCAAAAACAGGCTATGTGTGACATTAACCACTTCAGTACTG 960
QY 961 TTTATGCGCGACAGATTAATCTGCAAAATCTCGGCGCGCACTGAGCTCAACTGG 1020
Db 961 TTTATGCGCGACAGATTAATCTGCAAAATCTCGGCGCGCACTGAGCTCAACTGG 1020
QY 1021 ACCGTTCCCGGTGACCGGATTAACCGCGCGCGGCTGATGATGCTGTTGAACGCTGG 1080
Db 1021 ACCGTTCCCGGTGACCGGATTAACCGCGCGCGGCTGATGATGCTGTTGAACGCTGG 1080
QY 1081 CGTCGCTAAGCAGATTAACAGCCAGTGAATTCAGGTTTCGCTGCTTCCAGCTTTACAG 1140
Db 1081 CGTCGCTAAGCAGATTAACAGCCAGTGAATTCAGGTTTCGCTGCTTCCAGCTTTACAG 1140
QY 1141 CAGATGCTGATTAACCGCGCTGTCTCATTAATTAACCGCGCGGAGAGGTTAACTGACC 1200
Db 1141 CAGATGCTGATTAACCGCGCTGTCTCATTAATTAACCGCGCGGAGAGGTTAACTGACC 1200
QY 1201 CTGCGAGATGATGAAGGCAATGCGAGGCGATGTGCTGTTGCAAGGTTTAAAGCA 1260
Db 1201 CTGCGAGATGATGAAGGCAATGCGAGGCGATGTGCTGTTGCAAGGTTTAAAGCA 1260
QY 1261 ATCGTGAATGAAGCAGCATACCGGCGGTGAGTTGAGATCTCATC 1306
Db 1261 ATCGTGAATGAAGCAGCATACCGGCGGTGAGTTGAGATCTCATC 1306

RESULT 5
AR636185
LOCUS AR636185 1323 bp DNA linear PAT 20-APR-2005
DEFINITION Sequence 1 from patent US 6855365.
ACCESSION AR636185
VERSION AR636185.1 GI:62768868
KEYWORDS
SOURCE .
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1323)
AUTHORS Short,J.M., Kretz,K.A., Gray,K.A., Barton,N.R., Garrett,J.B.,
O'Donoghue,B. and Mathur,B.J.
TITLE Recombinant bacterial phylases and uses thereof
JOURNAL Patent: US 6855365-A 1 15-FEB-2005;
Diversa Corporation; San Diego, CA

FEATURES
source 1..1323
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 96.7%; Score 1265; DB 2; Length 1323;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 ATGAAGCGATCTTAATCCATTTTATCTCTTGATTCCTGTTAACCCCGCAATCGCA 60
 Db 1 ATGAAGCGATCTTAATCCATTTTATCTCTTGATTCCTGTTAACCCCGCAATCGCA 60
 QY 61 TTCGCTCAGAGTGAAGCCGAGCTGAAGCTGGAAGTGTGTGATTTGTCACTGCTCATG 120
 Db 61 TTCGCTCAGAGTGAAGCCGAGCTGAAGCTGGAAGTGTGTGATTTGTCACTGCTCATG 120
 QY 121 GTGCGTGTCCAACCAAGGCGCAACGCAATGATGCAAGGATGCAACCCCAACGATGCGCA 180
 Db 121 GTGCGTGTCCAACCAAGGCGCAACGCAATGATGCAAGGATGCAACCCCAACGATGCGCA 180
 QY 181 ACCTGGCCGGTAAACTGGGTGAGCTGACACCGCGGTGTGTGATGCTTAATGCGCTATCTC 240
 Db 181 ACCTGGCCGGTAAACTGGGTGAGCTGACACCGCGGTGTGTGATGCTTAATGCGCTATCTC 240
 QY 241 GGACATTAATGCGCTGACGCGTGTGTGAGCGCAACGCAATGCTGCTTAATGCTGCTGCGCG 300
 Db 241 GGACATTAATGCGCTGACGCGTGTGTGAGCGCAACGCAATGCTGCTGCTGCTGCGCG 300
 QY 301 CAGTCTGTGCAAGTGTGCGGATTAATGCTGATGTCAGACGATACCCGTTAAACAGCGCA 360
 Db 301 CAGTCTGTGCAAGTGTGCGGATTAATGCTGATGTCAGACGATACCCGTTAAACAGCGCA 360
 QY 361 GCGTCTGCGCGCGGCTGCGCACTGACCTGCACTGCAATACCGTACATACCGAGGCAATAG 420
 Db 361 GCGTCTGCGCGCGGCTGCGCACTGACCTGCACTGCAATACCGTACATACCGAGGCAATAG 420
 QY 421 TCCAGTCCCGATCCGTTATTTAATCTCTTAATAAAGTGGCTTGGCAATGATTAACGCG 480
 Db 421 TCCAGTCCCGATCCGTTATTTAATCTCTTAATAAAGTGGCTTGGCAATGATTAACGCG 480
 QY 481 AACGTGATGACGCGATCTCTGACAGGCGGAGGCTCAATGCTGATCTTAACCGGCGAT 540
 Db 481 AACGTGATGACGCGATCTCTGACAGGCGGAGGCTCAATGCTGATCTTAACCGGCGAT 540
 QY 541 TATCAACGCGCGTTCGCGCACTGGAAGGCGGCTTAATTTCCGCAATCAAACTTGTGCG 600
 Db 541 TATCAACGCGCGTTCGCGCACTGGAAGGCGGCTTAATTTCCGCAATCAAACTTGTGCG 600
 QY 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGACGCAATTAACATCGAATCTC 660
 Db 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGACGCAATTAACATCGAATCTC 660
 QY 661 AAGGTGAGCGCGGATGCTGTCTCAATTAACCGGTGCGGTAAAGCTGCGATCAATGCTGACG 720
 Db 661 AAGGTGAGCGCGGATGCTGTCTCAATTAACCGGTGCGGTAAAGCTGCGATCAATGCTGACG 720
 QY 721 GAGATATTTCTCTGCAACAGCAGAGGAAATGCGGAGCCGCGGTGGGGAAGATCAAC 780
 Db 721 GAGATATTTCTCTGCAACAGCAGAGGAAATGCGGAGCCGCGGTGGGGAAGATCAAC 780
 QY 781 GATTCAACAGTGAACAACCTTGGCTAAGTTGCAATACGCGCAATTTGATTGCTACAA 840
 Db 781 GATTCAACAGTGAACAACCTTGGCTAAGTTGCAATACGCGCAATTTGATTGCTACAA 840
 QY 841 CGCAGCGCAAGAGTTGCGCGACGCGCGCACCCGTTATTAAGATTGAATCAAGACAGCG 900
 Db 841 CGCAGCGCAAGAGTTGCGCGACGCGCGCACCCGTTATTAAGATTGAATCAAGACAGCG 900
 QY 901 TTGAAGCGCCCATCCACCGCAAAACAGGCGTATGCTGACATTAACCACTTCACTACG 960
 Db 901 TTGAAGCGCCCATCCACCGCAAAACAGGCGTATGCTGACATTAACCACTTCACTACG 960
 QY 961 TTTATCGCGGACAGATCTAATCTGGAATCTCGGCGGCGCACTGAGGCTCAACTG 1020
 Db 961 TTTATCGCGGACAGATCTAATCTGGAATCTCGGCGGCGCACTGAGGCTCAACTG 1020
 QY 1021 ACGCTTCCCGGTACGCGGATTAACGCGCGCGAGGTGTGAATCTGTGTTGAACGCTG 1080
 Db 1021 ACGCTTCCCGGTACGCGGATTAACGCGCGCGAGGTGTGAATCTGTGTTGAACGCTG 1080
 QY 1081 CGTGGCGTAAAGGATTAACGCGCGAGTTCAGGTTTGGCTGTCTTCAAGCTTACAG 1140

Db 1081 CGTGGCGTAAAGGATTAACGCGCGAGTTCAGGTTTGGCTGTCTTCAAGCTTACAG 1140
 QY 1141 CAGATGCGGTATTAACCGCGCTGTCTTAATTAACCGCGCGGAGGATGAATCTGAC 1200
 Db 1141 CAGATGCGGTATTAACCGCGCTGTCTTAATTAACCGCGCGGAGGATGAATCTGAC 1200
 QY 1201 CTGGAGGATGTAAGAGCGGAAATGCGAGGCGATGTTGCTGGCAGGTTTACGCA 1260
 Db 1201 CTGGAGGATGTAAGAGCGGAAATGCGAGGCGATGTTGCTGGCAGGTTTACGCA 1260
 QY 1261 ATCTGTAATGAAGACGCAATACCGCGCTGCACTTGAATCTCATC 1306
 Db 1261 ATCTGTAATGAAGACGCAATACCGCGCTGCACTTGAATCTCATC 1306

RESULT 6
 AX052713
 LOCUS AX052713 1323 bp DNA linear PART 12-JUN-2001
 DEFINITION Sequence 1 from Patent WO0071728.
 ACCESSION AX052713
 VERSION AX052713.1 GI:12226903
 KEYWORDS
 SOURCE
 ORGANISM
 Escherichia coli
 Escherichia coli
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.

REFERENCE
 1 Short, J.M. and Kretz, K.A.
 Recombinant bacterial phytases and uses thereof
 JOURNAL Patent: WO 0071728-A 1 30-NOV-2000;
 Diversa Corporation (US)
 Location/Qualifiers

FEATURES
 source
 1..1323
 /organism="Escherichia coli"
 /mol_type="unassigned DNA"
 /db_xref="taxon:562"
 1..1323
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC21810.1"
 /transl_table=1
 /db_xref="GI:12226904"

CDS
 1..1323
 /translation="MKATLIPFLSLILPVTQSAFQSEPEIKESVIVSRHGVAP
 TKATQMDVTPDAMPVPLGMLTPPGELIALVGYRQRLVADGLAKGCPDS
 GOVALIADVDERTKTRGEAPFAAGLAPDAIVHTQADSSDPLFNPKTGVCQADNA
 NVTDAIISRAGSIALDFGHRQTAFLRELRLVNPQSNLCKREKODESCSLTALPS
 ELKVSADNVSILGAVSILASMLTEI FILQDAGMBPGMGRITDSHOWMTLISINAP
 YLQRTPEVASRSRATPLDILMAALTPHPQKAYGVTLPTSVLPIAGHDTRNLNAG
 ALLENWTLPEQPDNTPREGELVFERWRRLSNSQWIOVSLVFOVLQMRDKTPUSLNT
 PPGEVKLTLACBERRNAQMCISLAGFTDIVEARI PACSLSRHHHHH"
 ORIGIN

Query Match 96.7%; Score 1265; DB 2; Length 1323;
 Beet Local Similarity 98.0%; Pred. No. 0;
 Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 ATGAAGCGATCTTAATCCATTTTATCTCTTGATTCCTGTTAACCCCGCAATCGCA 60
 Db 1 ATGAAGCGATCTTAATCCATTTTATCTCTTGATTCCTGTTAACCCCGCAATCGCA 60
 QY 61 TTCGCTCAGAGTGAAGCCGAGCTGAAGCTGGAAGTGTGTGATTTGTCACTGCTCATG 120
 Db 61 TTCGCTCAGAGTGAAGCCGAGCTGAAGCTGGAAGTGTGTGATTTGTCACTGCTCATG 120
 QY 121 GTGCGTGTCCAACCAAGGCGCAACGCAATGATGCAAGGATGCAACCCCAACGATGCGCA 180
 Db 121 GTGCGTGTCCAACCAAGGCGCAACGCAATGATGCAAGGATGCAACCCCAACGATGCGCA 180
 QY 181 ACCTGGCCGGTAAACTGGGTGAGCTGACACCGCGGTGTGTGATGCTTAATGCGCTATCTC 240
 Db 181 ACCTGGCCGGTAAACTGGGTGAGCTGACACCGCGGTGTGTGATGCTTAATGCGCTATCTC 240

QY 241 GGAATTATCTGGCGTCAAGCGTCTGTGTAACCGAGATTGCTCCTTAATATGTCCTCCG 300
DB 241 GGAATTATCTGGCGTCAAGCGTCTGTGTAACCGAGATTGCTGCGAATAAGGCTGCGCG 300
QY 301 CAGTCTGGTCAAGTCCGATTAATGCTGATGTCAGACGAGCGCTTACCCGTAACAGCGGAA 360
DB 301 CAGTCTGGTCAAGTCCGATTAATGCTGATGTCAGACGAGCGCTTACCCGTAACAGCGGAA 360
QY 361 GCCTTCGCGCGCGGCTGCGACCTGACCTGTCAATTAACCGTACATACCCAGGAGATACG 420
DB 361 GCCTTCGCGCGCGGCTGCGACCTGACCTGTCAATTAACCGTACATACCCAGGAGATACG 420
QY 421 TCCAGTCCCGATCCGTTATTTAATCTCTTAATAAATGCGGTTTCCAACTGATTAACGCG 480
DB 421 TCCAGTCCCGATCCGTTATTTAATCTCTTAATAAATGCGGTTTCCAACTGATTAACGCG 480
QY 481 AACGTGACTGAACGATCTCTGAGAGGCGAGAGGATTAATGCTGATTTACCGGGCAT 540
DB 481 AACGTGACTGAACGATCTCTGAGAGGCGAGAGGATTAATGCTGATTTACCGGGCAT 540
QY 541 TATCAACGCGGCTTCCGGAAGTGGAACTGTAATTTTCCGCAATCAAACTTGTGC 600
DB 541 CCGCAACGCGGCTTCCGGAAGTGGAACTGTAATTTTCCGCAATCAAACTTGTGC 600
QY 601 CTTAACGCTGAAGAAACAGACGAAAGCTTTCATTAACGACGATTAACATCGGAATC 660
DB 601 CTTAACGCTGAAGAAACAGACGAAAGCTTTCATTAACGACGATTAACATCGGAATC 660
QY 661 AAGGTGAGCGCGGCTGTGTCTCATTAACCGGTGCGGTGAACCTTCGATCAATGCTGACG 720
DB 661 AAGGTGAGCGCGGCTGTGTCTCATTAACCGGTGCGGTGAACCTTCGATCAATGCTGACG 720
QY 721 GAGATATTTCTCCGCAACAGACGAGGAAATGCGGAGCGGCGGAGGAGGAAAGATCAC 780
DB 721 GAGATATTTCTCCGCAACAGACGAGGAAATGCGGAGCGGCGGAGGAGGAAAGATCAC 780
QY 781 GATTCAACAGAGTGAACACCTTGTGTAATGTTGCAATACGCGCAATTTGATTTGCTACA 840
DB 781 GATTCAACAGAGTGAACACCTTGTGTAATGTTGCAATACGCGCAATTTGATTTGCTACA 840
QY 841 CGCACGCGAAGGTTGCGCGGAGCGCGGCAACCGGTTATTAATGATTTGATCAAGACGCG 900
DB 841 CGCACGCGAAGGTTGCGCGGAGCGCGGCAACCGGTTATTAATGATTTGATCAAGACGCG 900
QY 901 TTGACGCGCCATTCACCGCAAAACAGCGGTATGATTAACCACTTCAGTACTG 960
DB 901 TTGACGCGCCATTCACCGCAAAACAGCGGTATGATTAACCACTTCAGTACTG 960
QY 961 TTATTCGCGGACAGATTAATCTGCAATCTCGCGGCGCACTGAGCTCAACTG 1020
DB 961 TTATTCGCGGACAGATTAATCTGCAATCTCGCGGCGCACTGAGCTCAACTG 1020
QY 1021 ACGCTTCCCGGTCAAGCGGATTAACAGCGCGGAGGTGTAATGCTGTTGAAACGCTGCG 1080
DB 1021 ACGCTTCCCGGTCAAGCGGATTAACAGCGCGGAGGTGTAATGCTGTTGAAACGCTGCG 1080
QY 1081 CGTGGGCTTAAGCGATTAACAGCGAGTGAATTCAGGTTTCGCTGCTTCAGACTTTACAG 1140
DB 1081 CGTGGGCTTAAGCGATTAACAGCGAGTGAATTCAGGTTTCGCTGCTTCAGACTTTACAG 1140
QY 1141 CAGATGCGTGAATAAAGCGCGCTGTCTAATAATACGCGCGGAGAGGTGAATACTGACC 1200
DB 1141 CAGATGCGTGAATAAAGCGCGCTGTCTAATAATACGCGCGGAGAGGTGAATACTGACC 1200
QY 1201 CTGCGAGGATGTAAGAGCGAAATGCGAGGCGCATGTGCTGTTGCGAGGTTTAAACGAA 1260
DB 1201 CTGCGAGGATGTAAGAGCGAAATGCGAGGCGCATGTGCTGTTGCGAGGTTTAAACGAA 1260
QY 1261 ATCGGATGAAGACGATACCGGCGGTGCACTTTGATCTCATC 1306
DB 1261 ATCGGATGAAGACGATACCGGCGGTGCACTTTGATCTCATC 1306

RESULT 7
AX338966 1323 bp DNA linear PAT 09-JAN-2002
LOCUS
DEFINITION
Sequence 1 from Patent WO0189317.
ACCESSION
AX338966
VERSION
AX338966.1 GI:18129102
KEYWORDS
SOURCE
ORGANISM
Escherichia coli
Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE
1
AUTHORS
Short, J.M., Kretz, K.A. and O'Donoghue, B.
TITLE
Dietary aids and methods of use thereof
JOURNAL
Patent: WO 0189317-A 1 29-NOV-2001;
CORPORATION (US)
DIVERSA
FEATURES
source
Location/Qualifiers
1..1323
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"
1..1323
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAD20553.1"
/db_xref="GI:18129103"
CDS
1..1323
/translation="MKRLIPPLSLILPLIPPOSAPFAPQSEBELKESVIVIRBHVAP
TXATQMODVTPPAMPTWPKLGMTPRGEELIAYIGHYRQRLVADGLAKKCCPS
GVAITADVDERTKTEBAPAGLAPCATITVADTSPDLFPLFKTGVCOLDNA
NVDAILSRAGSIAFPTGHRQTFRELEVLNFPQSNCLKREKDESCSLQALPS
ELVSDVNSLITGAVSLASMLTEIPLIQKQMPBEGMGTITDSHQNLTLSLHNOF
YLQRTPEVRSRATPLDLIMAAITPHPOKQAVGVTLPSTVLAFLAGHDTNLNIG
ALFNMTLPGQNPRTPEGELVPERHWRRLSDNSQWLOVSLVFOTLOMRPKTPLSLNT
PPEVKKLTLAECERNAQKCSLAGTQIVNEARIRACSLRSHHHHH"
ORIGIN
Query Match 96.7%; Score 1265; DB 2; Length 1323;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 1 ATGAAAGCATCTTAATCCATTTTATCTCTTCTGATTCGCTTAACCCCGCAATCTGCA 60
DB 1 ATGAAAGCATCTTAATCCATTTTATCTCTTCTGATTCGCTTAACCCCGCAATCTGCA 60
QY 61 TTGCTCAGAGTGAACCGGAGCTGAAGCTGGAAGGTGATGATTCAGTCCGATGCT 120
DB 61 TTGCTCAGAGTGAACCGGAGCTGAAGCTGGAAGGTGATGATTCAGTCCGATGCT 120
QY 121 GTGCGTCTCAACCAAGGCGACGACATGATGAGAGTCAACCCAGACGATGCGCA 180
DB 121 GTGCGTCTCAACCAAGGCGACGACATGATGAGAGTCAACCCAGACGATGCGCA 180
QY 181 ACCTGCGGTTAAACTGGGTGAGCTGACACCGCGCGGTGTGATTAATCGCTATCTC 240
DB 181 ACCTGCGGTTAAACTGGGTGAGCTGACACCGCGCGGTGTGATTAATCGCTATCTC 240
QY 241 GGAATTATCTGGCGTCAAGCGTCTGTGTAACCGGAGTTGCTGCTTAATATGCTGCGCG 300
DB 241 GGAATTATCTGGCGTCAAGCGTCTGTGTAACCGGAGTTGCTGCTTAATATGCTGCGCG 300
QY 301 CAGTCTGGTCAAGTCCGATTAATGCTGATGTCAGACGAGCGCTTACCCGTAACAGCGGAA 360
DB 301 CAGTCTGGTCAAGTCCGATTAATGCTGATGTCAGACGAGCGCTTACCCGTAACAGCGGAA 360
QY 361 GCCTTCGCGCGCGGCTGCGACCTGACCTGTCAATTAACCGTACATACCCAGGAGATACG 420
DB 361 GCCTTCGCGCGCGGCTGCGACCTGACCTGTCAATTAACCGTACATACCCAGGAGATACG 420
QY 421 TCCAGTCCCGATCCGTTATTTAATCTCTTAATAAATGCGGTTTCCAACTGATTAACGCG 480
DB 421 TCCAGTCCCGATCCGTTATTTAATCTCTTAATAAATGCGGTTTCCAACTGATTAACGCG 480

QY 481 AACGTGACTGACGCGATCTCTCGAGAGGAGGAGGATTCATTTGCTGACTTTACCGGGCAT 540
 DB 481 AACGTGACTGACGCGATCTCTCGAGAGGAGGAGGATTCATTTGCTGACTTTACCGGGCAT 540
 QY 541 TATCAAAACGGCGTTTCGCGAACTGGAAACGGGTGCTTAATTTTCGCAATCAAACTTGTGC 600
 DB 541 CGGCAAAACGGCGTTTCGCGAACTGGAAACGGGTGCTTAATTTTCGCAATCAAACTTGTGC 600
 QY 601 CTTAAACGTGAAGAAACAGAACGAAAGCTTTCTATTACGAGGATTTACCATTCGAACTC 660
 DB 601 CTTAAACGTGAAGAAACAGAACGAAAGCTTTCTATTACGAGGATTTACCATTCGAACTC 660
 QY 661 AAGGTAGCGCGACTGTCTCATTTAACCGGTGCGGTAAAGCTGCGATCAATGCGAGC 720
 DB 661 AAGGTAGCGCGACTGTCTCATTTAACCGGTGCGGTAAAGCTGCGATCAATGCGAGC 720
 QY 721 GAGATATTTCTCTGCAACAAACAGAGGAAATGCGGAGCCGCGGTGCGAAAGATCAC 780
 DB 721 GAGATATTTCTCTGCAACAAACAGAGGAAATGCGGAGCCGCGGTGCGAAAGATCAC 780
 QY 781 GATTCAACACAGTGAACACCTTGCTAATGTTGCTAATACGCGCAATTTGATTGCTACAA 840
 DB 781 GATTCAACACAGTGAACACCTTGCTAATGTTGCTAATACGCGCAATTTGATTGCTACAA 840
 QY 841 CGCAGCGCGAGAGTTGCCCGCAGCGCGCGCCGTTATTTAGATTTGATCAAGACAGC 900
 DB 841 CGCAGCGCGAGAGTTGCCCGCAGCGCGCGCCGTTATTTAGATTTGATCAAGACAGC 900
 QY 901 TTGAGCGCCCATCAACCGCAAAAACAGCGGTATGCTGATTAACCATTCAGTGTG 960
 DB 901 TTGAGCGCCCATCAACCGCAAAAACAGCGGTATGCTGATTAACCATTCAGTGTG 960
 QY 961 TTTATCGCGGACACAGATATCTAATCTGCGAAATCTCGCGCGCGCATCTGAGTCAACTG 1020
 DB 961 TTTATCGCGGACACAGATATCTAATCTGCGAAATCTCGCGCGCGCATCTGAGTCAACTG 1020
 QY 1021 ACGCTTCCCGGTGACCGCGGATTAACGCGCGCGAGGTGTAATCTGTTTGAAGCTG 1080
 DB 1021 ACGCTTCCCGGTGACCGCGGATTAACGCGCGCGAGGTGTAATCTGTTTGAAGCTG 1080
 QY 1081 GGTGCGCTTAAGCGATTAACGCGAGTGTGATTCAGTTCGTTGCTTCCAGACTTACG 1140
 DB 1081 GGTGCGCTTAAGCGATTAACGCGAGTGTGATTCAGTTCGTTGCTTCCAGACTTACG 1140
 QY 1141 CAGATCGGTATTAACCGCGCTGTCTAATTAATACGCGCGCGAGAGTGAATCTGACC 1200
 DB 1141 CAGATCGGTATTAACCGCGCTGTCTAATTAATACGCGCGCGAGAGTGAATCTGACC 1200
 QY 1201 CTGGCAGGATGTAAGAGGGAATGCGAGGCGATGTGTTGTTGGCAGGTTTACGCA 1260
 DB 1201 CTGGCAGGATGTAAGAGGGAATGCGAGGCGATGTGTTGTTGGCAGGTTTACGCA 1260
 QY 1261 ATCGTGAATGAAGACGATACCGGCGTGCAGTTCATTC 1306
 DB 1261 ATCGTGAATGAAGACGATACCGGCGTGCAGTTCATTC 1306

RESULT 8
 AX356566 1323 bp DNA linear PAT 06-FEB-2002
 LOCUS AX356566
 DEFINITION Sequence 1 from Patent WO0190333.
 ACCESSION AX356566
 VERSION AX356566.1 GI:18621053
 KEYWORDS
 ORGANISM
 Escherichia coli
 Escherichia coli
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 Enterobacteriaceae; Escherichia.
 REFERENCES
 1 Short, J.M., Kretz, K.A., Gray, K.A., Barton, N.R., Garret, J.B. and
 O'Donoghue, E.
 Recombinant bacterial phytases and uses thereof
 JOURNAL Patent: WO 0190333-A 1 29-NOV-2001;

FEATURES
 source
 CDS
 DIVERSA CORPORATION (US)
 location/Qualifiers
 1..1323
 /organism="Escherichia coli"
 /mol_type="unassigned DNA"
 /db_xref="taxon:562"
 1..1323
 /note="unassigned protein product"
 /codon_start=1
 /transl_table=11
 /protein_id="CAD23025.1"
 /db_xref="GI:18621054"
 /translation="MKAILIPILSLIPLPQSAFPOSEPLKLESLVIVRHGVAP
 TKATQIMQDVTVPDAMPPTPVLCGLMTPPRGGLIATVIGYQORLVADGLAKGCPQS
 QGVATLADVDERIKRTGSAFAGALPDCATLVHVDADTSSDPFLNPLKTVCCQDNA
 NVTDAILSRAGSIALDPTGHRQTAFLRELEKVLNFPQSLCKREKQDSGLVQALPS
 ELKVSADNVSLTGAVLSMLTBIFLIQAOAGMEPGMGRIITDSHQMTLLSLHNAOF
 YLQORTPEVARSRAATPLDLIMAAITLPHPOKQAVGTVLPSVLPIAGHDTLAVLGG
 ALBINMTLRQDPDPTPRGELVPRWRRLSDNSQWIOVSLPQTLQMRDKTPLSLNT
 PPGRYKLTLLACBRNMQMCSLAFQIYVNEARLPACSLASHHHHH"

ORIGIN
 Query Match 96.7%; Score 1265; DB 2; Length 1323;
 Best Local Similarity 98.0%; Pred. No. 0;
 Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 ATGAAAGGATCTTAATCCATTTTATCTCTGATTCGTTAACCCTGGATCTGCA 60
 DB 1 ATGAAAGGATCTTAATCCATTTTATCTCTGATTCGTTAACCCTGGATCTGCA 60
 QY 61 TTCGCTCAGATGAGACCGGAGCTGAAGCTGAAAGTGTGTGATGTCAGTGTCTATGTT 120
 DB 61 TTCGCTCAGATGAGACCGGAGCTGAAGCTGAAAGTGTGTGATGTCAGTGTCTATGTT 120
 QY 61 TTCGCTCAGATGAGACCGGAGCTGAAGCTGAAAGTGTGTGATGTCAGTGTCTATGTT 120
 DB 61 TTCGCTCAGATGAGACCGGAGCTGAAGCTGAAAGTGTGTGATGTCAGTGTCTATGTT 120
 QY 121 GTGCGTGTCCCAACCAAGCCACGCACTGATGACAGATGTCAACCCAGACGATGGCA 180
 DB 121 GTGCGTGTCCCAACCAAGCCACGCACTGATGACAGATGTCAACCCAGACGATGGCA 180
 QY 181 ACTTGGCCGGTAAACTGGGTGAGCTGACACCGCGCGGTGTGAGCTAATGCGCTATCTC 240
 DB 181 ACTTGGCCGGTAAACTGGGTGAGCTGACACCGCGCGGTGTGAGCTAATGCGCTATCTC 240
 QY 181 ACTTGGCCGGTAAACTGGGTGAGCTGACACCGCGCGGTGTGAGCTAATGCGCTATCTC 240
 DB 181 ACTTGGCCGGTAAACTGGGTGAGCTGACACCGCGCGGTGTGAGCTAATGCGCTATCTC 240
 QY 241 GGAATTAATGCGCGCTGACAGCTGTGTGACCGACGATGCTGCTTAATGTGTCGCGCCG 300
 DB 241 GGAATTAATGCGCGCTGACAGCTGTGTGACCGACGATGCTGCTTAATGTGTCGCGCCG 300
 QY 241 GGAATTAATGCGCGCTGACAGCTGTGTGACCGACGATGCTGCTTAATGTGTCGCGCCG 300
 DB 241 GGAATTAATGCGCGCTGACAGCTGTGTGACCGACGATGCTGCTTAATGTGTCGCGCCG 300
 QY 301 CAGTCTGAGTGCAGTGCAGATTAATGCTGATGTGACGAGCGGTAACCGTAAACAGCGCA 360
 DB 301 CAGTCTGAGTGCAGTGCAGATTAATGCTGATGTGACGAGCGGTAACCGTAAACAGCGCA 360
 QY 361 GCTTTCGCGCGCGGCTGGCACTGACTGTCGCAATTAACCGTAAACAGCGCA 360
 DB 361 GCTTTCGCGCGCGGCTGGCACTGACTGTCGCAATTAACCGTAAACAGCGCA 360
 QY 421 TCAGTCCCGCATCCGCTTAATTTTAAATCTGCAAACTGCGCTTTCGCAACTGAACTGC 480
 DB 421 TCAGTCCCGCATCCGCTTAATTTTAAATCTGCAAACTGCGCTTTCGCAACTGAACTGC 480
 QY 481 AACGTGACTGACGCGATCTCTGAGAGGCGAGAGGATTCATTTGCTGACTTTACCGGGCAT 540
 DB 481 AACGTGACTGACGCGATCTCTGAGAGGCGAGAGGATTCATTTGCTGACTTTACCGGGCAT 540
 QY 541 TATCAAAACGGCGTTTCGCGAACTGGAAACGGGTGCTTAATTTTCGCAATCAAACTTGTGC 600
 DB 541 TATCAAAACGGCGTTTCGCGAACTGGAAACGGGTGCTTAATTTTCGCAATCAAACTTGTGC 600
 QY 601 CTTAAACGTGAAGAAACAGAACGAAAGCTTTCTATTACGAGGATTTACCATTCGAACTC 660
 DB 601 CTTAAACGTGAAGAAACAGAACGAAAGCTTTCTATTACGAGGATTTACCATTCGAACTC 660
 QY 661 AAGGTAGCGCGACTGTGTCTCATTTAACCGGTGCGGTAAAGCTGCGATCAATGCGAGC 720
 DB 661 AAGGTAGCGCGACTGTGTCTCATTTAACCGGTGCGGTAAAGCTGCGATCAATGCGAGC 720

QY	721	GAGATATTTCCTCTGCACAAGCACAGGGAATGCCGAGGCCGGGGTGGGGAAGATCAACC	780
Db	721	GAGATATTTCCTCTGCACAAGCACAGGGAATGCCGAGGCCGGGGTGGGGAAGATCAACC	780
QY	781	GATTCAACACAGGTGGACAACCTGTCTAGTTTGCATAACGGCGCAATTGATTTCGTACAA	840
Db	781	GATTCAACACAGGTGGACAACCTGTCTAGTTTGCATAACGGCGCAATTGATTTCGTACAA	840
QY	841	CGCAGCGCAGAGGTTGCCCGCAGCCGCCCAACCCCGTATTAGATTGATCAAGACAGCG	900
Db	841	CGCAGCGCAGAGGTTGCCCGCAGCCGCCCAACCCCGTATTAGATTGATCAAGACAGCG	900
QY	901	TTGACAGCCCATCCACCGCAAAAACAGCCGTAATGATGACATTACCACTTCAGTGTG	960
Db	901	TTGACAGCCCATCCACCGCAAAAACAGCCGTAATGATGACATTACCACTTCAGTGTG	960
QY	961	TTTATCGCGCGGACAGATCTAATCTGGCAATCTCGCGCGGCGCATGAGACTCAACTGG	1020
Db	961	TTTATCGCGCGGACAGATCTAATCTGGCAATCTCGCGCGGCGCATGAGACTCAACTGG	1020
QY	1021	AACGCTTCCCGGTCAACCGCATPAACAACGCCCGCAGGTGTGTGAATCTGTTTGAACGCTGG	1080
Db	1021	AACGCTTCCCGGTCAACCGCATPAACAACGCCCGCAGGTGTGTGAATCTGTTTGAACGCTGG	1080
QY	1081	CGTCCGGCTAAGCGATPAACAAGCAGTGGATTCAAGTTTGGCTGTCTTCCAGACTTTACAG	1140
Db	1081	CGTCCGGCTAAGCGATPAACAAGCAGTGGATTCAAGTTTGGCTGTCTTCCAGACTTTACAG	1140
QY	1141	CAGATGCGGTGATPAAAACGCCGCTGCATTTAATAAGCCGCCCGCAGAGAGATGAACCTGACC	1200
Db	1141	CAGATGCGGTGATPAAAACGCCGCTGCATTTAATAAGCCGCCCGCAGAGAGATGAACCTGACC	1200
QY	1201	CTGGCAGGATGTGAAGAGCGAAATGCCAGGGCAGTGTTCGTTGGCAGGTTTACGCA	1260
Db	1201	CTGGCAGGATGTGAAGAGCGAAATGCCAGGGCAGTGTTCGTTGGCAGGTTTACGCA	1260
QY	1261	ATCGTGATGAAGACAGCATACCGCGCTGCAGTTTGAGATCTCATC 1306	
Db	1261	ATCGTGATGAAGACAGCATACCGCGCTGCAGTTTGAGATCTCATC 1306	

RESULT 9	
U00096_10	
WPCOMMENT	
Sequence split into 47 fragments	LOCUS U00096 Accession U00096
Fragment Name	Begin End
U00096_00	1 110000
U00096_01	100001 210000
U00096_02	200001 310000
U00096_03	300001 410000
U00096_04	400001 510000
U00096_05	500001 610000
U00096_06	600001 710000
U00096_07	700001 810000
U00096_08	800001 910000
U00096_09	900001 1010000
U00096_10	1000001 1100000
U00096_11	1100001 1210000
U00096_12	1200001 1310000
U00096_13	1300001 1410000
U00096_14	1400001 1510000
U00096_15	1500001 1610000
U00096_16	1600001 1710000
U00096_17	1700001 1810000
U00096_18	1800001 1910000
U00096_19	1900001 2010000
U00096_20	2000001 2110000
U00096_21	2100001 2210000
U00096_22	2200001 2310000
U00096_23	2300001 2410000
U00096_24	2400001 2510000
U00096_25	2500001 2610000

Query	Match	Best Local Similarity	Matches 1278; Conservative	96.7%; 98.3%;	Score 1264.8; Pred. No. 0;	DB 15;	Length 110000;			
U00096_26	260001	271000								
U00096_27	270001	281000								
U00096_28	280001	291000								
U00096_29	290001	301000								
U00096_30	300001	311000								
U00096_31	310001	321000								
U00096_32	320001	331000								
U00096_33	330001	341000								
U00096_34	340001	351000								
U00096_35	350001	361000								
U00096_36	360001	371000								
U00096_37	370001	381000								
U00096_38	380001	391000								
U00096_39	390001	401000								
U00096_40	400001	410000								
U00096_41	410001	421000								
U00096_42	420001	431000								
U00096_43	430001	441000								
U00096_44	440001	451000								
U00096_45	450001	461000								
U00096_46	460001	463675								
Continuation (11 of 47) of U00096 Becherichia coli K-12 MG1655										
Query Match	96.7%;	Score 1264.8;	DB 15;	Length 110000;						
Best Local Similarity	98.3%;	Pred. No. 0;								
Matches 1278; Conservative	0;	Mismatches 22;	Indels 0;	Gaps 0;						
1	ATGAAAGCATCTTAATCCATTTTATCTCTTGATTCGGTTAACCCGCAATCTGCA	60								
39840	ATGAAAGCATCTTAATCCATTTTATCTCTTGATTCGGTTAACCCGCAATCTGCA	39899								
61	TTGCTCAGAGTAGAGCCGAGACTGAAGCTGAAAGTGTGATGTCAGTCTCATGCT	120								
39900	TTGCTCAGAGTAGAGCCGAGACTGAAGCTGAAAGTGTGATGTCAGTCTCATGCT	39959								
121	GTGCTGTCTTCAACCAAGCCACGCACTGATGAGATGTCACCCGAGCGCATGCGCA	180								
39960	GTGCTGTCTTCAACCAAGCCACGCACTGATGAGATGTCACCCGAGCGCATGCGCA	40019								
181	ACCTGGCCGGTAAACCTGGGTAGCTGACACCGCGCGGTGAGCTAATCGCCTATCTC	240								
40020	ACCTGGCCGGTAAACCTGGGTAGCTGACACCGCGCGGTGAGCTAATCGCCTATCTC	40079								
241	GGACATTAATGAGCGCTGCTGTAGCCGACGGAATGCTGCTTAAATGTGCTGCCG	300								
40080	GGACATTAATGAGCGCTGCTGTAGCCGACGGAATGCTGCTTAAATGTGCTGCCG	40139								
301	CAGTCTGTGACGATCGCGATTATTGCTGATGTGACGAGCGTACCCTTAAACAGCGCA	360								
40140	CAGTCTGTGACGATCGCGATTATTGCTGATGTGACGAGCGTACCCTTAAACAGCGCA	40199								
361	GGCTTCGCGCGGGGTGGACCTGACTGTGCAATAACCGTACATACCAGCAGATACG	420								
40200	GGCTTCGCGCGGGGTGGACCTGACTGTGCAATAACCGTACATACCAGCAGATACG	40259								
421	TCGAGTCCCGATCCGTTATTAAATCTCTTAAACCTGCGCTTGCATCTGATACCGG	480								
40260	TCGAGTCCCGATCCGTTATTAAATCTCTTAAACCTGCGCTTGCATCTGATACCGG	40319								
481	AACGTGACTGACGCGATCTCTGAGAGGGCAGAGGGTCAATTGCTGACTTTACCGG	540								
40320	AACGTGACTGACGCGATCTCTGAGAGGGCAGAGGGTCAATTGCTGACTTTACCGG	40379								
541	TATCAACCGCGCTTGGCGCACTGGAACGGGTGCTTAATTTCCGCAATCAACTGTGC	600								
40380	CGGCAACCGCGCTTGGCGCACTGGAACGGGTGCTTAATTTCCGCAATCAACTGTGC	40439								
601	CTTAAACGTGAGAAACAGACGAAAGCTGTTCATTAAACGACGATTAACATCGAACTC	660								
40440	CTTAAACGTGAGAAACAGACGAAAGCTGTTCATTAAACGACGATTAACATCGAACTC	40499								
6										

Db 40500 AAGTGAGCGCCGACAAATGCTCTATTAAACCGGTGCGGTAAAGCTCGCATCAATGCTGAGC 40559
Qy 721 GAGATATTTCTCTGCAACAAGACAGGGAAATCCCGAGCCGGGTGGGAAAGATCAACC 780
Db 40560 GAGATATTTCTCTGCAACAAGACAGGGAAATCCCGAGCCGGGTGGGAAAGATCAACC 40619
Qy 781 GATTCAACACAGTGGAAACAATTGTCTAATGTCATATGACCGCAATTTGATTGCTACAA 840
Db 40620 GATTCAACACAGTGGAAACAATTGTCTAATGTCATATGACCGCAATTTGATTGCTACAA 40679
Qy 841 CGCAGCGCCAGAGTGTGCGCGGACCGCGCCACCCCGTATTGATTTGATCAAGACAGCG 900
Db 40680 CGCAGCGCCAGAGTGTGCGCGGACCGCGCCACCCCGTATTGATTTGATCAAGACAGCG 40739
Qy 901 TTGAGCGCCCATTCACCCGAAAAACAAGGGTATGCTGTATCAATTAACCCATTCAAGTCTG 960
Db 40740 TTGAGCGCCCATTCACCCGAAAAACAAGGGTATGCTGTATCAATTAACCCATTCAAGTCTG 40799
Qy 961 TTATTCGCGGACACAGATATCTATCTGCAAAATCTCGCGGCGACCTGAGCTCAACTGG 1020
Db 40800 TTATTCGCGGACACAGATATCTATCTGCAAAATCTCGCGGCGACCTGAGCTCAACTGG 40859
Qy 1021 ACGCTTCCCGGTCAAGCCGGAATAACAGCCGCGCAAGTGTGAATCTGTGTTTGAACGCTGG 1080
Db 40860 ACGCTTCCCGGTCAAGCCGGAATAACAGCCGCGCAAGTGTGAATCTGTGTTTGAACGCTGG 40919
Qy 1081 CGTCCGCTTAAGCATTAACAGCCAGTGTGATTCAGGTTTCTGTCTTCCAGCTTTACG 1140
Db 40920 CGTCCGCTTAAGCATTAACAGCCAGTGTGATTCAGGTTTCTGTCTTCCAGCTTTACG 40979
Qy 1141 CAGATCGGTATTAACGCGCGTGTCAATTAATACGCGCGCGGAGGATGAACCTGACC 1200
Db 40980 CAGATCGGTATTAACGCGCGTGTCAATTAATACGCGCGCGGAGGATGAACCTGACC 41039
Qy 1201 CTGCGAGATGTGAAGAGCGAAATGCGAGGCGCATGTGTTCTGTGCGAGGTTTACGCA 1260
Db 41040 CTGCGAGATGTGAAGAGCGAAATGCGAGGCGCATGTGTTCTGTGCGAGGTTTACGCA 41099
Qy 1261 ATCCGTAATGAAGCAGCATACCGCGCGTGCATTTGAGAT 1300
Db 41100 ATCCGTAATGAAGCAGCATACCGCGCGTGCATTTGAGAT 41139

RESULT 10

AP009048_10

WFOCOMMENT

Sequence split into 47 fragments LOCUS AP009048 Accession AP009048

Fragment Name	Begin	End
AP009048_00	1	110000
AP009048_01	100001	210000
AP009048_02	200001	310000
AP009048_03	300001	410000
AP009048_04	400001	510000
AP009048_05	500001	610000
AP009048_06	600001	710000
AP009048_07	700001	810000
AP009048_08	800001	910000
AP009048_09	900001	1010000
AP009048_10	1000001	1110000
AP009048_11	1100001	1210000
AP009048_12	1200001	1310000
AP009048_13	1300001	1410000
AP009048_14	1400001	1510000
AP009048_15	1500001	1610000
AP009048_16	1600001	1710000
AP009048_17	1700001	1810000
AP009048_18	1800001	1910000
AP009048_19	1900001	2010000
AP009048_20	2000001	2110000
AP009048_21	2100001	2210000
AP009048_22	2200001	2310000
AP009048_23	2300001	2410000
AP009048_24	2400001	2510000

AP009048_25 2500001 2610000
AP009048_26 2600001 2710000
AP009048_27 2700001 2810000
AP009048_28 2800001 2910000
AP009048_29 2900001 3010000
AP009048_30 3000001 3110000
AP009048_31 3100001 3210000
AP009048_32 3200001 3310000
AP009048_33 3300001 3410000
AP009048_34 3400001 3510000
AP009048_35 3500001 3610000
AP009048_36 3600001 3710000
AP009048_37 3700001 3810000
AP009048_38 3800001 3910000
AP009048_39 3900001 4010000
AP009048_40 4000001 4110000
AP009048_41 4100001 4210000
AP009048_42 4200001 4310000
AP009048_43 4300001 4410000
AP009048_44 4400001 4510000
AP009048_45 4500001 4610000
AP009048_46 4600001 4646332

Continuation (11 of 47) of AP009048 from base 1000001 (AP009048 Escherichia coli W3110 D

Query Match 96.7%; Score 1264.8; DB 15; Length 110000;
Best Local Similarity 98.3%; Pred. No. 0; Mismatches 22; Indels 0; Gaps 0;
Matches 1278; Conservative 0

Qy 1 ATGAAGGAGATCTTAATCCCAATTTTATCTCTGATTCGTTAACCCTGCAATCTGCA 60
Db 41039 ATGAAGGAGATCTTAATCCCAATTTTATCTCTGATTCGTTAACCCTGCAATCTGCA 41098
Qy 61 TTCCGCTCAAGTAGAGCCGAGAGCTGAAGCTGGAAGTGTGTGATGTGTCATGTCTATGCT 120
Db 41099 TTCCGCTCAAGTAGAGCCGAGAGCTGAAGCTGGAAGTGTGTGATGTGTCATGTCTATGCT 41158
Qy 121 GTGCGTGTCCAAACCAAGCCAGCACTGATGACAGATGTCACCCAGAGCATGGCCA 180
Db 41159 GTGCGTGTCCAAACCAAGCCAGCACTGATGACAGATGTCACCCAGAGCATGGCCA 41218
Qy 181 ACTTGGCCGCTTAACCTGGTGAAGCTGACACCGCGCGTGTGAGCTAATGCGCTATCTC 240
Db 41219 ACTTGGCCGCTTAACCTGGTGAAGCTGACACCGCGCGTGTGAGCTAATGCGCTATCTC 41278
Qy 241 GGAACATTAATGAGCGCTGACGCTGTGTGAGCCGACGAGATTGCTGCTTAATGTGCTGCCG 300
Db 41279 GGAACATTAATGAGCGCTGACGCTGTGTGAGCCGACGAGATTGCTGCTGCTTAATGTGCTGCCG 41338
Qy 301 CAGTCTGTCAAGTGTGCGCATTTATGCTGATGTCACAGAGCGTTAATCCCTGTAACAGGCCGA 360
Db 41339 CAGTCTGTCAAGTGTGCGCATTTATGCTGATGTCACAGAGCGTTAATCCCTGTAACAGGCCGA 41398
Qy 361 GCGTTGCGCGCGGAGCTGAGCACTGATGTCGTAATTAACCTGTAACCTGAGGAGATGACG 420
Db 41399 GCGTTGCGCGCGGAGCTGAGCACTGATGTCGTAATTAACCTGTAACCTGAGGAGATGACG 41458
Qy 421 TCCAGTCCGATCCGTTATTTATCTCTTAATAAACTGCGCTTGGCACTGAGATACGCG 480
Db 41459 TCCAGTCCGATCCGTTATTTATCTCTTAATAAACTGCGCTTGGCACTGAGATACGCG 41518
Qy 481 AACGTGATGACGGAATCTCTGAGAGGCGAGAGGCTTAATGCTGATTTTACCGGGCAT 540
Db 41519 AACGTGATGACGGAATCTCTGAGAGGCGAGAGGCTTAATGCTGATTTTACCGGGCAT 41578
Qy 541 TATCAAAAGCGGCTTTCGGAAGTGAAGCGGAGCTTAATTTTCGCAATCAAACTTGTGC 600
Db 41579 TATCAAAAGCGGCTTTCGGAAGTGAAGCGGAGCTTAATTTTCGCAATCAAACTTGTGC 41638
Qy 601 CTTAAAGCTGAAGAACAGAGCAAGAGCTTTCATTAACGAGGATTAACATCGGAATC 660
Db 41639 CTTAAAGCTGAAGAACAGAGCAAGAGCTTTCATTAACGAGGATTAACATCGGAATC 41698
Qy 661 AAGGTAGCGCGGACGTGTCTCATTAACCGGTGCGTAACCTCGCATCAATGCTGACG 720

|||||
Db 41699 AAGGTGAGCGCGCAATGCTCAATTAACCGGTGGGTAAAGCTTCGACATCAATGCTGACG 41758
Qy 721 GAGATATTTCTCCGCAACAAGCAACGGAATGCCGAGCCGGGGTGGGAAAGATCAAC 780
Db 41759 GAGATATTTCTCCGCAACAAGCAACGGAATGCCGAGCCGGGGTGGGAAAGATCAAC 41818
Qy 781 GATTCAACACAGTGAACACCTTGCTAAGTTTGCAATAACGCAATTTGCTACAA 840
Db 41819 GATTCAACACAGTGAACACCTTGCTAAGTTTGCAATAACGCAATTTGCTACAA 41878
Qy 841 CGCAGCGCAGAGTTGCGCGCAGCCGCGCAACCCCGTTATTAGATTGATCAAGACGCG 900
Db 41879 CGCAGCGCAGAGTTGCGCGCAGCCGCGCAACCCCGTTATTAGATTGATCAAGACGCG 41938
Qy 901 TTGAGCGCCCATCAACGCAAAAAACAGCGATAGTGTGACATTAACCACTTCAGTGTG 960
Db 41939 TTGAGCGCCCATCAACGCAAAAAACAGCGATAGTGTGACATTAACCACTTCAGTGTG 41998
Qy 961 TTATATCGCGGACACGATACTAATCTGGCAATCTCGCGCGCACTGGAAGCTCAACTGG 1020
Db 41999 TTATATCGCGGACACGATACTAATCTGGCAATCTCGCGCGCACTGGAAGCTCAACTGG 42058
Qy 1021 ACGCTTCCGGGTCAAGCCGATTAACAAGCGCCAGGTGTGAACCTGGTTTGAACGCTGG 1080
Db 42059 ACGCTTCCGGGTCAAGCCGATTAACAAGCGCCAGGTGTGAACCTGGTTTGAACGCTGG 42118
Qy 1081 CGTGGCGTAAGCGATTAACAGCAGCAGTGTGATTCAGTTTGGCTTCCAACTTTACAG 1140
Db 42119 CGTGGCGTAAGCGATTAACAGCAGCAGTGTGATTCAGTTTGGCTTCCAACTTTACAG 42178
Qy 1141 CAGATGCGTGTATTAACACCGCTGTCAATTAATAACGCGCCGAGAGAGTGAACCTGACC 1200
Db 42179 CAGATGCGTGTATTAACACCGCTGTCAATTAATAACGCGCCGAGAGAGTGAACCTGACC 42238
Qy 1201 CTGGCAGGATGTGAAGACGCAATATCGCAGGCGCATGTGTTCTTGGCAGAGTTTAAACGCA 1260
Db 42239 CTGGCAGGATGTGAAGACGCAATATCGCAGGCGCATGTGTTCTTGGCAGAGTTTAAACGCA 42298
Qy 1261 ATCGTGATTAAGACGCAATACCGCGCTGTCATTAATAACGCGCCGAGAGAGTGAACCTGACC 1300
Db 42299 ATCGTGATTAAGACGCAATACCGCGCTGTCATTAATAACGCGCGTGTGTAAT 42338

RESULT 11
AR636189
LOCUS AR636189 1901 bp DNA linear PAT 20-APR-2005
DEFINITION Sequence 6 from patent US 6855365.
ACCESSION AR636189
VERSION AR636189.1 GI:62768872
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1901)
AUTHORS Short,J.M., Kretz,K.A., Gray,K.A., Barton,N.R., Garretl,J.B.,
O'Donoghue,E. and Mathur,E.J.
TITLE Recombinant bacterial phycases and uses thereof
JOURNAL Patent: US 6855365-A 6 15-FEB-2005;
Diversa Corporation; San Diego, CA
FEATURES
source 1..1901
/organism="Unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 96.6%; Score 1263.8; DB 2; Length 1901;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 1277; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 1 ATGAAAGGATCTTAATCCATTTTATCTCTTGATTCGGTTAACCCGCGCAATCTGCA 60
Db 188 ATGAAAGGATCTTAATCCATTTTATCTCTTGATTCGGTTAACCCGCGCAATCTGCA 247

Qy 61 TTGCTCAGAGTGAAGCCGAGACTGAAGCTGAAAAGTGTGTGATTGTCAGTTCGTATGCT 120
Db 248 TTGCTCAGAGTGAAGCCGAGACTGAAGCTGAAAAGTGTGTGATTGTCAGTTCGTATGCT 307
Qy 121 GTGCTGTCTCCAAACCAAGGCCACGCAACTGATGCAAGATGTCAACCCAGACGATGGCCA 180
Db 308 GTGCTGTCTCCAAACCAAGGCCACGCAACTGATGCAAGATGTCAACCCAGACGATGGCCA 367
Qy 181 ACCGTGCGGGTAAACCTGGGTGAGCTGACACCGCGGGTGTGAGCTAAATCGCTATCTC 240
Db 368 ACCGTGCGGGTAAACCTGGGTGAGCTGACACCGCGGGTGTGAGCTAAATCGCTATCTC 427
Qy 241 GGAATTAATCTGGCGTCAAGCTCTGTGTAGCCGACGCAATTTGCTTAATGTGCTGCGCG 300
Db 428 GGAATTAATCTGGCGCAAGCGTCTGTGTAGCCGACGCAATTTGCTGGGAAAAAGGCTGCGCG 487
Qy 301 CAGTCTGTCAAGTCCGCAATTAATTTCTGATGTGACGAGCTAACCCGTAAACAAGCGCAA 360
Db 488 CAGTCTGTCAAGTCCGCAATTAATTTCTGATGTGACGAGCTAACCCGTAAACAAGCGCAA 547
Qy 361 GCCTTCCGCGCGCGCTGGGCACTGACTGTGCAATTAACGCTAATACCGCAGGAGATACG 420
Db 548 GCCTTCCGCGCGCGCTGGGCACTGACTGTGCAATTAACGCTAATACCGCAGGAGATACG 607
Qy 421 TCCAGTCCCGATCGTTATTTAAATCCTCTAATAAATGCGCTTTGSCAACTGATAACGCG 480
Db 608 TCCAGTCCCGATCGTTATTTAAATCCTCTAATAAATGCGCTTTGSCAACTGATAACGCG 667
Qy 481 AACGTGACTGACGCGATCTCTGAGAGGGCAGAGGGCTCAATTCGTGACTTTACCGGGCAT 540
Db 668 AACGTGACTGACGCGATCTCTGAGAGGGCAGAGGGCTCAATTCGTGACTTTACCGGGCAT 727
Qy 541 TATCAACCGCGCTTTCGCGCAACTGGAACGGGTCTTAATTTTCGCGCAATGAACCTGTGC 600
Db 728 CCGCAACCGCGCTTTCGCGCAACTGGAACGGGTCTTAATTTTCGCGCAATGAACCTGTGC 787
Qy 601 CTTTAAAGTGAAGAAACAGACGAAAGCTGTTCATTAACGACGCAATTACCATCGAACTC 660
Db 788 CTTTAAAGTGAAGAAACAGACGAAAGCTGTTCATTAACGACGCAATTACCATCGAACTC 847
Qy 661 AAGGTGAGCGCGCACTGTGTCTCATTAACCGGTGCGGTAAAGCTTCGATCAATGCTGACG 720
Db 848 AAGGTGAGCGCGCACTGTGTCTCATTAACCGGTGCGGTAAAGCTTCGATCAATGCTGACG 907
Qy 721 GAGATATTTCTCTGCAACAAGCAACAGGAATGCGGAGCGGGGGTGGGAAAGATACCC 780
Db 908 GAGATATTTCTCTGCAACAAGCAACAGGAATGCGGAGCGGGGGTGGGAAAGATACCC 967
Qy 781 GATTCAACACAGTGAACACCTTGCTAAGTTTGATGAATACGCGCAATTGATTTGCTACAA 840
Db 968 GATTCAACACAGTGAACACCTTGCTAAGTTTGATGAATACGCGCAATTGATTTGCTACAA 1027
Qy 841 CGCAGCGCAAGAGTTGCGCGCAGCCGCGCAACCCCGTTATTAGATTGATCAAGACGCG 900
Db 1028 CGCAGCGCAAGAGTTGCGCGCAGCCGCGCAACCCCGTTATTAGATTGATCAAGACGCG 1087
Qy 901 TTGAGCGCCCATCAACGCAAAAAACAGGCGTATGTGTGACATTAACCACTTCAGTGTG 960
Db 1088 TTGAGCGCCCATCAACGCAAAAAACAGGCGTATGTGTGACATTAACCACTTCAGTGTG 1147
Qy 961 TTATATGCGCGGACACGATACTAATCTGGCAATCTCGCGCGCGCACTGAGAGCTCAACTGG 1020
Db 1148 TTATATGCGCGGACACGATACTAATCTGGCAATCTCGCGCGCGCACTGAGAGCTCAACTGG 1207
Qy 1021 ACGCTTCCGGGTCAAGCCGATTAACAAGCGCCAGGTGTGAACCTGTGTTGAACGCTGG 1080
Db 1208 ACGCTTCCGGGTCAAGCCGATTAACAAGCGCCAGGTGTGAACCTGTGTTGAACGCTGG 1267
Qy 1081 CGTGGCTTAAGCGATTAACAGCAGGTGATTCAGGTTTCGTGTCTTCGACATTTACAG 1140
Db 1268 CGTGGCTTAAGCGATTAACAGCAGGTGATTCAGGTTTCGTGTCTTCGACATTTACAG 1327

Qy 1141 CAGATGCGTGAATAAACCGCGCTGTCATTAAATACCGCCGCGAGAGGTGAACCTGACC 1200
Db 1328 CAGATGCGTGAATAAACCGCGCTGTCATTAAATACCGCCGCGAGAGGTGAACCTGACC 1387
Qy 1201 CTGGCAGAGTGTGAAGAGCGAAATGCGAGCGGCGATGTGTTCTGGCAGGTTTAAAGCAA 1260
Db 1388 CTGGCAGAGTGTGAAGAGCGAAATGCGAGCGGCGATGTGTTCTGGCAGGTTTAAAGCAA 1447
Qy 1261 ATCGTAATGAAGACGCGATACCGGCGTGCAGTTTGAGAT 1300
Db 1448 ATCGTAATGAAGACGCGATACCGGCGTGCAGTTTGAT 1487

RESULT 12
AR636190
LOCUS AR636190 1901 bp DNA linear PAT 20-APR-2005
DEFINITION Sequence 7 from patent US 685365.
ACCESSION AR636190
VERSION AR636190.1 GI:62768873
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE
AUTHORS 1 (bases 1 to 1901)
TITLE Short J.M., Kretz, K.A., Gray, K.A., Barton, N.R., Garrett, J.B.,
O'Donoghue, E. and Mathur, E.J.
JOURNAL Recombinant bacterial phylases and uses thereof
Diverse Corporation; San Diego, CA
LOCATION/Qualifiers
FEATURES
source 1..1901
/organism="Unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 96.6%; Score 1263.8; DB 2; Length 1901;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 1277; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 1 ATGAAAGGATCTTAATCCCATTTTATCTCTTGATTCGCTTAACCCCGCAATCTGCA 60
Db 188 ATGAAAGGATCTTAATCCCATTTTATCTCTTGATTCGCTTAACCCCGCAATCTGCA 247
Qy 61 TTGCGTCAAGTGAACCGGAGCTGAAGCTGAAGTGTGATGTTGTCACTGATGATGAT 120
Db 248 TTGCGTCAAGTGAACCGGAGCTGAAGTGTGATGTTGTCACTGATGATGATGATGAT 307
Qy 121 GTGCGTGTCCAAACGAGGCGCAAGCACTGATGAGATGTCAACCCCGCAAGGATGCGA 180
Db 308 GTGCGTGTCCAAACGAGGCGCAAGCACTGATGAGATGTCAACCCCGCAAGGATGCGA 367
Qy 181 ACTGCGCGGTAAACCTGGGTGAGTGAACCGCGGCTGTGAGCTAATGCTATCTC 240
Db 368 ACTGCGCGGTAAACCTGGGTGAGTGAACCGCGGCTGTGAGCTAATGCTATCTC 427
Qy 241 GGACATTAATGCGGCTGACGCTGTGATGCGGATGCTGCTTAATGTGAGTGGCCG 300
Db 428 GGACATTAATGCGGCTGACGCTGTGATGCGGATGCTGCTTAATGTGAGTGGCCG 487
Qy 301 CAGTGTGATGAGTGTGAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 360
Db 488 CAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 547
Qy 361 GCTTTCGCGCGCGGCTGCGCACTGATGCTGCAATTAACGTAATCCGAGCAGATACG 420
Db 548 GCTTTCGCGCGCGGCTGCGCACTGATGCTGCAATTAACGTAATCCGAGCAGATACG 607
Qy 421 TCGAGTCCGATCCGCTTAATTAATCTCTTAATAATGCGGTTTGCACTGATTAACGCG 480
Db 608 TCGAGTCCGATCCGCTTAATTAATCTCTTAATAATGCGGTTTGCACTGATTAACGCG 667
Qy 481 AACGTAAGTGAACGCGATCTCGAGAGGCGAGGAGGTCAATTTGCTTAACGCGGAT 540

Db 668 AACGTAAGTGAACGCGATCTCGAGAGGCGAGGAGGTCAATTTGCTTAACCGGCGAT 727
Qy 541 TATCAACGCGGCTTTCGCAACTGGAACGCGGTCTTAATTTTCGCAATCAACTTGTGC 600
Db 728 CGGCAACGCGGCTTTCGCAACTGGAACGCGGTCTTAATTTTCGCAATCAACTTGTGC 787
Qy 601 CTTAAACGTGAAGAACAGAGCGAAGCTGTTCATTAAAGCGGCGATTCACATCGGAATCC 660
Db 788 CTTAAACGTGAAGAACAGAGCGAAGCTGTTCATTAAAGCGGCGATTCACATCGGAATCC 847
Qy 661 AAGTGAACGCGCGATGCTGTCTATTAAACCGTGTGCGTAAGCTGTGATCAATGCTGACG 720
Db 848 AAGTGAACGCGCGCGATGCTGTCTATTAAACCGTGTGCGTAAGCTGTGATCAATGCTGACG 907
Qy 781 GATTCAACAGTGAAGAACCTTGTCTAATGCTTAACCGGCAATTTGATTTGCTACAA 840
Db 968 GATTCAACAGTGAAGAACCTTGTCTAATGCTTAACCGGCAATTTGATTTGCTACAA 1027
Qy 841 CGCAGCGAGAGTGTGCGCGAGCGCGCGCAACCCGTTATTAGATTTGATCAAGCAGCG 900
Db 1028 CGCAGCGAGAGTGTGCGCGAGCGCGCGCGCAACCCGTTATTAGATTTGATCAAGCAGCG 1087
Qy 901 TTGAGCGCCCATTCACCGCAAAACAGGCGTATGTGTGAACATTACCACTTCACTGCTG 960
Db 1088 TTGAGCGCCCATTCACCGCAAAACAGGCGTATGTGTGAACATTACCACTTCACTGCTG 1147
Qy 961 TTTATGCGCGGAGCAACGATTAATTTGCAAAATCTCGCGCGCGCATCTGAGTCACTG 1020
Db 1148 TTTATGCGCGGAGCAACGATTAATTTGCAAAATCTCGCGCGCGCATCTGAGTCACTG 1207
Qy 1021 ACGCTTCCCGGTGACCGCGATTAACAGCGCGCGAGGTGTGAACCTGTGTAAGCGCTG 1080
Db 1208 ACGCTTCCCGGTGACCGCGATTAACAGCGCGCGAGGTGTGAACCTGTGTAAGCGCTG 1267
Qy 1081 CGTGGCTTAAGCGATTAACAGCGCGATGATTTGAGTTTGTCTTTCAGACTTTACG 1140
Db 1268 CGTGGCTTAAGCGATTAACAGCGCGATGATTTGAGTTTGTCTTTCAGACTTTACG 1327
Qy 1141 CAGATGCGTGAATAAACCGCGCTGTCTTAATTAATACCGCCGCGAGAGGTGAACCTGACC 1200
Db 1328 CAGATGCGTGAATAAACCGCGCTGTCTTAATTAATACCGCCGCGAGAGGTGAACCTGACC 1387
Qy 1201 CTGGCAGAGTGTGAAGAGCGAAATGCGAGCGGCGATGTGTTGGCAGGTTTAAAGCAA 1260
Db 1388 CTGGCAGAGTGTGAAGAGCGAAATGCGAGCGGCGATGTGTTGGCAGGTTTAAAGCAA 1447
Qy 1261 ATCGTAATGAAGACGCGATACCGGCGTGCAGTTTGAGAT 1300
Db 1448 ATCGTAATGAAGACGCGATACCGGCGTGCAGTTTGAT 1487

RESULT 13
AX356571
LOCUS AX356571 1901 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 6 from Patent WO0190333.
ACCESSION AX356571
VERSION AX356571.1 GI:18621058
KEYWORDS
SOURCE
ORGANISM
Escherichia coli
Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE
AUTHORS 1
TITLE Short J.M., Kretz, K.A., Gray, K.A., Barton, N.R., Garrett, J.B. and
JOURNAL Recombinant bacterial phylases and uses thereof
DIVERSA CORPORATION (US)
LOCATION/Qualifiers

SOURCE

1. .1901
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"

ORIGIN

```

Query Match      96.6%; Score 1263.8; DB 2; Length 1901;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 1277; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY      1 ATGAAGCGATCTTAATCCATTTTATCTCTTGATTCGGTTAACCCGCAATCTGCA 60
DB      188 ATGAAGCGATCTTAATCCATTTTATCTCTTGATTCGGTTAACCCGCAATCTGCA 247
QY      61 TTGCGTCAAGTGAAGCCGAGCTGAAGCTGAAGTGGATGTCAGTCAGTCAGT 120
DB      248 TTGCGTCAAGTGAAGCCGAGCTGAAGTGGATGTCAGTCAGTCAGT 307
QY      121 GTGCGTCTCCAAACCAAGGCGCAACTGATGAGATGTCACCCGAGAGCATGGCCA 180
DB      308 GTGCGTCTCCAAACCAAGGCGCAACTGATGAGATGTCACCCGAGAGCATGGCCA 367
QY      181 ACCTGGCCGGTAAACTGAGTGAAGCTGACACCGCGGGTGTGAGCTAATGCTTATCTC 240
DB      368 ACCTGGCCGGTAAACTGAGTGAAGCTGACACCGCGGGTGTGAGCTAATGCTTATCTC 427
QY      241 GGAGATTACTGGCGCTGAGCGCTGTGTAGCGAGATTGCTGCTTAATGTGGTCCGCG 300
DB      428 GGAGATTACTGGCGCGAGCGCTGTGTAGCGAGATTGCTGCGAAGAAAGGCGTCCGCG 487
QY      301 CAGTCTGTGTCAGGTGCGCAATTAATGCTGATGTCAGACAGCGTACCCGTTAAACAGGCGAA 360
DB      488 CAGTCTGTGTCAGGTGCGCAATTAATGCTGATGTCAGACAGCGTACCCGTTAAACAGGCGAA 547
QY      361 GCCTTCGCGCGCGGCTGCGCACTGATCTGTGCAATACCGTACATACCAGGCGAGTACG 420
DB      548 GCCTTCGCGCGCGGCTGCGCACTGATCTGTGCAATACCGTACATACCAGGCGAGTACG 607
QY      421 TCCAGTCCCGATCGGTTATTAATCTCTTAATAACGTGGCGTTGGCAATGAGTAAAGCGG 480
DB      608 TCCAGTCCCGATCGGTTATTAATCTCTTAATAACGTGGCGTTGGCAATGAGTAAAGCGG 667
QY      481 AACGTGACTGAACGCGATCTCTGAGAGGCGAGAGGGTCAATGTGCACTTTTACCGGGCAT 540
DB      668 AACGTGACTGAACGCGATCTCTGAGAGGCGAGAGGGTCAATGTGCACTTTTACCGGGCAT 727
QY      541 TATCAAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACCTGTGCG 600
DB      728 CCGCAAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACCTGTGCG 787
QY      601 CTTAAACGTGAGAAACAGGAAGGAAAGCTTCAATTAACGAGGCAATTAACATCGGAATCTC 660
DB      788 CTTAAACGTGAGAAACAGGAAGGAAAGCTTCAATTAACGAGGCAATTAACATCGGAATCTC 847
QY      661 AAGGTGAGCGCGCATGTGTCTCATTTAACCGGTGCGGTAGACCTTCGATCAATGCTGACG 720
DB      848 AAGGTGAGCGCGCATGTGTCTCATTTAACCGGTGCGGTAGACCTTCGATCAATGCTGACG 907
QY      721 GAGATATTTCTCTGCAACAGCACAGGAAATGCCGAGCCGGGGTGGGAAAGATCAC 780
DB      908 GAGATATTTCTCTGCAACAGCACAGGAAATGCCGAGCCGGGGTGGGAAAGATCAC 967
QY      781 GATTCAACACGAGTGAACACCTGTAGTTTGAATTAACGGGCAATTTGATTTGCTACAA 840
DB      968 GATTCAACACGAGTGAACACCTGTAGTTTGAATTAACGGGCAATTTGATTTGCTACAA 1027
QY      841 CGCAGCGCAGAGGTTCGCGAGCGCGCACCCCGTATTAGATTGATCAAGACAGCG 900
DB      1028 CGCAGCGCAGAGGTTCGCGAGCGCGCACCCCGTATTAGATTGATCAAGACAGCG 1087
QY      901 TTGAGCGCCCATCCACCGCAAAAACAGGCGTATGTTGATTAACCATTCAGTGTG 960
DB      1088 TTGAGCGCCCATCCACCGCAAAAACAGGCGTATGTTGATTAACCATTCAGTGTG 1147

```

```

QY      961 TTTATCGCCGGACACGATACTAATCTGGCAATCTCGGCGGCGCACTGAGCTCAACTGG 1020
DB      1148 TTTATCGCCGGACACGATACTAATCTGGCAATCTCGGCGGCGCACTGAGCTCAACTGG 1207
QY      1021 ACGCTTCCCGGTGACCGGATTAACACGCGCGCAGGTGTGAAGTGTGTTGAACGCTGG 1080
DB      1208 ACGCTTCCCGGTGACCGGATTAACACGCGCGCAGGTGTGAAGTGTGTTGAACGCTGG 1267
QY      1081 CGTCGGTAAAGCATTAACACGCGAGTGAATTCAGTTTCCGTGCTTCCAGACTTTACAG 1140
DB      1268 CGTCGGTAAAGCATTAACACGCGAGTGAATTCAGTTTCCGTGCTTCCAGACTTTACAG 1327
QY      1141 CAGATGCGTATTAACACGCGCTGTCAATTAATACCGCGCGGAGAGTGAACCTGACC 1200
DB      1328 CAGATGCGTATTAACACGCGCTGTCAATTAATACCGCGCGGAGAGTGAACCTGACC 1387
QY      1201 CTGGCAGATGTGAAGAGCGAAATGCGCAGGCGCATGTGTTGCTGGCAGGTTTACGCAA 1260
DB      1388 CTGGCAGATGTGAAGAGCGAAATGCGCAGGCGCATGTGTTGCTGGCAGGTTTACGCAA 1447
QY      1261 ATCGTGAATGAACACGCAATACCGGCGGTGCAATTGAGAT 1300
DB      1448 ATCGTGAATGAACACGCAATACCGGCGGTGCAATTGAGAT 1487

```

RESULT 14

```

AX356572      1901 bp      DNA      linear      PAT 06-FEB-2002
LOCUS
DEFINITION
Sequence 7 from Patent WO0190333.
ACCESSION
AX356572.1 GI:18621059
VERSION
KEYWORDS
SOURCE
ORGANISM

```

Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.

REFERENCE

1 Short, J.M., Kretz, K.A., Gray, K.A., Barton, N.R., Garrett, J.B. and O'Donoghue, E.
Recombinant bacterial phytases and uses thereof
Patent: WO 0190333-A 7 29-NOV-2001;
DIVERSA CORPORATION (US)

FEATURES

```

location/Qualifiers
source
1. .1901
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"

```

ORIGIN

```

Query Match      96.6%; Score 1263.8; DB 2; Length 1901;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 1277; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY      1 ATGAAGCGATCTTAATCCATTTTATCTCTTGATTCGGTTAACCCGCAATCTGCA 60
DB      188 ATGAAGCGATCTTAATCCATTTTATCTCTTGATTCGGTTAACCCGCAATCTGCA 247
QY      61 TTGCGTCAAGTGAAGCCGAGCTGAAGCTGAAGTGGATGTCAGTCAGTCAGT 120
DB      248 TTGCGTCAAGTGAAGCCGAGCTGAAGTGGATGTCAGTCAGTCAGT 307
QY      121 GTGCGTCTCCAAACCAAGGCGCAACTGATGAGATGTCACCCGAGAGCATGGCCA 180
DB      308 GTGCGTCTCCAAACCAAGGCGCAACTGATGAGATGTCACCCGAGAGCATGGCCA 367
QY      181 ACCTGGCCGGTAAACTGAGTGAAGCTGACACCGCGGGTGTGAGCTAATGCTTATCTC 240
DB      368 ACCTGGCCGGTAAACTGAGTGAAGCTGACACCGCGGGTGTGAGCTAATGCTTATCTC 427
QY      241 GGAGATTACTGGCGCTGAGCGTGTGTAGCGGAGATTGCTGCTTAATGTGGTGGCGG 300
DB      428 GGAGATTACTGGCGCTGAGCGTGTGTAGCGGAGATTGCTGCGAAGAAAGGCGTGGCGG 487

```

QY 301 CAGTCTGGTCAGAGTCGCGATTAATGTCGATGCGAGCGGTAACCGGTAACAGGCGGA 360
DB 488 CAGTCTGGTCAGAGTCGCGATTAATGTCGATGCGAGCGGTAACCGGTAACAGGCGGA 547
QY 361 GCGTTTCGCGCGGCGGCTGACCTGACCTGACGATTAACCGGTAACCGGTAACAGGCGGA 420
DB 548 GCGTTTCGCGCGGCGGCTGACCTGACCTGACGATTAACCGGTAACCGGTAACAGGCGGA 607
QY 421 TCCAGTCCCGATCCGTTATTTATCTCTTAATAAATCGGCGTTGCCAATGATTAACGCG 480
DB 608 TCCAGTCCCGATCCGTTATTTATCTCTTAATAAATCGGCGTTGCCAATGATTAACGCG 667
QY 481 AACGTAAGTGAAGGATCTGAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
DB 668 AACGTAAGTGAAGGATCTGAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 727
QY 541 TATCAAAACGCGCTTTCGCGAATGGAACGCGGCTTAAATTTCCGCAATCAACTTGTGC 600
DB 728 CCGCAAAACGCGCTTTCGCGAATGGAACGCGGCTTAAATTTCCGCAATCAACTTGTGC 787
QY 601 CTTAAACGTAAGAAACAGGAGGAAAGCTGTTCAATTAACGAGGCAATTAACGAGAACTC 660
DB 788 CTTAAACGTAAGAAACAGGAGGAAAGCTGTTCAATTAACGAGGCAATTAACGAGAACTC 847
QY 661 AAGGTAGAGGCGGCACTGTCTCATTAACCGGTCGGTAAAGGCTGCAATGCTGACG 720
DB 848 AAGGTAGAGGCGGCACTGTCTCATTAACCGGTCGGTAAAGGCTGCAATGCTGACG 907
QY 721 GAGATATTTCTCTCTCAACAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
DB 908 GAGATATTTCTCTCTCAACAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 967
QY 781 GATTCAACACAGTGAACAACCTTCTGTAAGTTGCAATACGCGCAATTTGATTGCTACAA 840
DB 968 GATTCAACACAGTGAACAACCTTCTGTAAGTTGCAATACGCGCAATTTGATTGCTACAA 1027
QY 841 CGCACGCGAGAGGTTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
DB 1028 CGCACGCGAGAGGTTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1087
QY 901 TTGACGCGGCAATCAACCGCAAAACAGGCGGTAAGTGTGACCAATTAACCACTTCACTGCTG 960
DB 1088 TTGACGCGGCAATCAACCGCAAAACAGGCGGTAAGTGTGACCAATTAACCACTTCACTGCTG 1147
QY 961 TTTATCGCGGCAACGATTAATCTGAGCAATCTCGCGGCGGCACTGAGCTCAACTGCG 1020
DB 1148 TTTATCGCGGCAACGATTAATCTGAGCAATCTCGCGGCGGCACTGAGCTCAACTGCG 1207
QY 1021 ACGGTTCCCGGTCAGCGGATTAACAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
DB 1208 ACGGTTCCCGGTCAGCGGATTAACAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1267
QY 1081 CGTGGCTTAAGCGATTAACAGCGCAAGTGAATTCAGTTTCGCTGCTTCCAGACTTACAG 1140
DB 1268 CGTGGCTTAAGCGATTAACAGCGCAAGTGAATTCAGTTTCGCTGCTTCCAGACTTACAG 1327
QY 1141 CAGATGCGTGAATAAAGCGCGCTGTCAATTAATACGCGCGCGGAGAGGTGAACCTGACC 1200
DB 1328 CAGATGCGTGAATAAAGCGCGCTGTCAATTAATACGCGCGCGGAGAGGTGAACCTGACC 1387
QY 1201 CTGGCAGATGAAGAGAGGAAATGCGAGGCGGCAATGTTGTTGCGGCGGTTTACGCA 1260
DB 1388 CTGGCAGATGAAGAGAGGAAATGCGAGGCGGCAATGTTGTTGCGGCGGTTTACGCA 1447
QY 1261 ATCGTAATGAAGACGCAATCCGCGCTGACAGTTTGAAGAT 1300
DB 1448 ATCGTAATGAAGACGCAATCCGCGCTGACAGTTTGAAT 1487

RESULT 15
ECOAPPA
LOCUS ECOAPPA 1901 bp DNA linear BCT 04-APR-2002

DEFINITION Escherichia coli periplasmic phosphoanhydride phosphohydrolase
(appa) gene, complete cds.
ACCESSION M58708
VERSION M58708.1 GI:145283
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCES
AUTHORS
TITLE
JOURNAL
PUBMED
FEATURES
source
location/Qualifiers
1. 1901
/organism="Escherichia coli"
/mol_type="genomic DNA"
/db_xref="taxon:562"
3. 1522
/gene="appa"
3. 8
/gene="appa"
16. 108
/gene="appa"
/note="ORF"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAV2085.1"
/db_xref="GI:145284"
/translation="MYILMFVGIILMCSITLVLVWLDPRLS"
118. 149
/gene="appa"
133. 138
/gene="appa"
156. 161
/gene="appa"
172. 177
/gene="appa"
188. 1486
/gene="appa"
/standard_name="acid phosphatase"
/codon_start=1
/transl_table=1
/product="periplasmic phosphoanhydride phosphohydrolase"
/protein_id="AAV2086.1"
/db_xref="GI:145285"
/translation="MKAILIPFLSLIPLPQSAFAQSEPELKLSEVIVSRHGRAP
TKATQMODVTPDAMPVPLGWLTPRGSELAYLGHYGRGLVADGLAKGCPQS
GOVALIADVERETKTEGAPAGAPCAITVHTQADTSPDPLFNLKTVGCOLDA
NYDAIISRAAGSIADPTGHRORTAFRELVANPQSNLCIKRKQDESGSTALPS
ELKXSADNVSLTGAVSLASLWLTIRFLAQACQMPREMGKITTSHQMTLISLANAP
YLQRTPEVARSATPULDLIKTALVTPPOKAYGVTLPTSVLFIAGHDTNLNLDG
ALBELMTLPGQDPTPGGELVFERWRLSDNSQIOWSLVFTLQOMRDTPLSLMT
PPGEVKTLLAGCEBRNAQGCISLAGTQIVNEARIIPACSL"
188. 253
/gene="appa"

sig_peptide

```
mat_peptide      254..1483
                  /gene="appa"
variation         /product="periplasmic phosphoanhydride phosphohydrolase"
                  299..301
                  /gene="appa"
                  /standard_name="H303A mutant"
                  /note="created by site directed mutagenesis"
                  /citation=[2]
                  299..301
variation         /gene="appa"
                  /standard_name="R16A mutant"
                  /note="created by site directed mutagenesis"
                  /citation=[2]
                  299..301
variation         /gene="appa"
                  /standard_name="R20A mutant"
                  /note="created by site directed mutagenesis"
                  /citation=[2]
                  299..301
variation         /gene="appa"
                  /standard_name="R92A mutant"
                  /note="created by site directed mutagenesis"
                  /citation=[2]
                  302..304
variation         /gene="appa"
                  /standard_name="H17N mutant"
                  /note="created by site directed mutagenesis"
                  /citation=[2]
                  1163..1165
variation         /gene="appa"
                  /standard_name="D304A mutant"
                  /note="created by site directed mutagenesis"
                  /citation=[3]
                  /phenotype="Aep 304 changed to Ala"
                  /replace="gct"
                  1163..1165
variation         /gene="appa"
                  /standard_name="D304E mutant"
                  /note="created by site directed mutagenesis"
                  /citation=[3]
                  /phenotype="Aep 304 changed to Glu"
                  /replace="gag"
                  1497..1522
terminator       /gene="appa"
                  /note="rho-independent terminator"
ORIGIN
Query Match      96.6%; Score 1263.8; DB 15; Length 1901;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 1277; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
QY      1 ATGAAAGCGATTTAATCCCTTTTATCTTCTTGATTCGGTTAAACCCCGCAATCTGCA 60
DB      188 ATGAAAGCGATTTAATCCCTTTTATCTTCTTGATTCGGTTAAACCCCGCAATCTGCA 247
QY      61 TTGCTCAGAGTGAAGCGGAGCTGAAGCTGGAAGTGTGATTTGATTCAGTCTGATAGT 120
DB      248 TTGCTCAGAGTGAAGCGGAGCTGAAGCTGGAAGTGTGATTTGATTCAGTCTGATAGT 307
QY      121 GTGCGTGTCCCAACCAAGGCCAGCAACTGATGACAGTGTACCCCGACAGCGCATGGCCA 180
DB      308 GTGCGTGTCCCAACCAAGGCCAGCAACTGATGACAGTGTACCCCGACAGCGCATGGCCA 367
QY      181 ACCTGGCCGGTAAACTGGGTGAGCTGACACCGCGCGTGTGATGCTTAATGCTTATCTC 240
DB      368 ACCTGGCCGGTAAACTGGGTGAGCTGACACCGCGCGTGTGATGCTTAATGCTTATCTC 427
QY      241 GGAATTACTGCGCGTCAAGCGCTGTGGTAGCGAGATTGCTGCTAAATGTGAGCTGCGCG 300
DB      428 GGAATTACTGCGCGTCAAGCGCTGTGGTAGCGAGATTGCTGCGGAAAAAGGCTGCGCG 487
QY      301 CAGTCTGTGACGCTGCGGATTAATGCTGATGTGACGAGCGTACCCGTTAAACAGCGGAA 360
|||||
```

```
DB      488 CAGTCTGTGACGCTGCGGATTAATGCTGATGTGACGAGCGTACCCGTTAAACAGCGGAA 547
QY      361 GCCTTGCGCGCGGGCTGAGCACTGACTGTGCAATAACCTTAATACCCGAGCATACG 420
DB      548 GCCTTGCGCGCGGGCTGAGCACTGACTGTGCAATAACCTTAATACCCGAGCATACG 607
QY      421 TCCAGTCCCGATCGTTAATTAATTCCTTAATAAATGAGCTTTTGCCAACTGATAACGCG 480
DB      608 TCCAGTCCCGATCGTTAATTAATTCCTTAATAAATGAGCTTTTGCCAACTGATAACGCG 667
QY      481 AACGTGACTGACCGGATCTCTGAGAGGGCAGAGGGTCAATTTGCTGACCTTAACCGGGCAT 540
DB      668 AACGTGACTGACCGGATCTCTGAGAGGGCAGAGGGTCAATTTGCTGACCTTAACCGGGCAT 727
QY      541 TATCAAAACGGCGTTTGCGCGAACTGGAACGGGTCTTAATTTTCCGCAATCAACTGTGC 600
DB      728 CCGCAACCGCGTTTGCGCGAACTGGAACGGGTCTTAATTTTCCGCAATCAACTGTGC 787
QY      601 CTTAAACGTGAGAAACAGACAGAAAGCTGTTCAATTAACGAGGCAATTACATCGSAATC 660
DB      788 CTTAAACGTGAGAAACAGACAGAAAGCTGTTCAATTAACGAGGCAATTACATCGSAATC 847
QY      661 AAGGTAGCGCGCACTGTGTCTCATTAACCGGTGCGGTAAAGCTTCGATCAATGCTGACG 720
DB      848 AAGGTAGCGCGCGCACTGTGTCTCATTAACCGGTGCGGTAAAGCTTCGATCAATGCTGACG 907
QY      721 GAGATATTTCTCTGCAACAAGCAAGGGAAATGCCGAGCGGGGGTGGGAAAGGATCACG 780
DB      908 GAGATATTTCTCTGCAACAAGCAAGGGAAATGCCGAGCGGGGGTGGGAAAGGATCACG 967
QY      781 GATTTCACACAGTGAACAACCTTGCTTAAGTTTGATTAACCGGCAATTTGATTTGCTACAA 840
DB      968 GATTTCACACAGTGAACAACCTTGCTTAAGTTTGATTAACCGGCAATTTGATTTGCTACAA 1027
QY      841 CGCACGCCAGAGGTTGCCCGCAGCGCGCAACCCCGTTAATTAAGTTTGAATCAAGACGCG 900
DB      1028 CGCACGCCAGAGGTTGCCCGCAGCGCGCAACCCCGTTAATTAAGTTTGAATCAAGACGCG 1087
QY      901 TTGAACGCCCATCAACCGGAAACAGGGCGTATGTGTGACATTAACCATTCAGTGTG 960
DB      1088 TTGAACGCCCATCAACCGGAAACAGGGCGTATGTGTGACATTAACCATTCAGTGTG 1147
QY      961 TTTATCGCCGAGACAGATTAATCTGGAATCTCGGCGCGCACTGAGCTCAACTGCG 1020
DB      1148 TTTATCGCCGAGACAGATTAATCTGGAATCTCGGCGCGCACTGAGCTCAACTGCG 1207
QY      1021 ACGCTTCCGGTCAACCGGATTAACAGCGCGCAGGTGTGAATCTGTTTGAACGCTCG 1080
DB      1208 ACGCTTCCGGTCAACCGGATTAACAGCGCGCAGGTGTGAATCTGTTTGAACGCTCG 1267
QY      1081 CGTCGGCTAAGCGTAAACAGCGAGTGAATCAAGTTTGCTGAGCTTCCAGACTTTACAG 1140
DB      1268 CGTCGGCTAAGCGTAAACAGCGAGTGAATCAAGTTTGCTGAGCTTCCAGACTTTACAG 1327
QY      1141 CAGATGCGTATAAAGCGCGCTGTCTTAATAACGCGCGCGAGAGGTGAACCTGACC 1200
DB      1328 CAGATGCGTATAAAGCGCGCTGTCTTAATAACGCGCGCGAGAGGTGAACCTGACC 1387
QY      1201 CTGGCAGGATGTGAAGAGCAAAATGCGCAGGCGATGTGTTGTTGGAGGTTTACGCAA 1260
DB      1388 CTGGCAGGATGTGAAGAGCAAAATGCGCAGGCGATGTGTTGTTGGAGGTTTACGCAA 1447
QY      1261 ATGTGAATGAAGACGATACCGGGGTGACGTTTGAGAT 1300
DB      1448 ATGTGAATGAAGACGATACCGGGGTGACGTTTGAGAT 1487
```

Search completed: June 14, 2006, 14:10:05
Job time : 7319.05 secs

Thu Jun 15 11:39:02 2006

us-10-601-319-9.rnt

Page 1

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocelleraction Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2006, 10:04:50 ; Search time 260.506 Seconds

(without alignments)
9394.823 Million cell updates/sec

Title: US-10-601-319-9

Perfect score: 1308
Sequence: 1 atgaagcgatcttaacc.....gcattcgatctcatcta 1308Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 93554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/ina/1_COMB.seq:*
- 2: /EMC_Celerra_SIDS3/ptodata/2/ina/5_COMB.seq:*
- 3: /EMC_Celerra_SIDS3/ptodata/2/ina/6A_COMB.seq:*
- 4: /EMC_Celerra_SIDS3/ptodata/2/ina/6B_COMB.seq:*
- 5: /EMC_Celerra_SIDS3/ptodata/2/ina/7_COMB.seq:*
- 6: /EMC_Celerra_SIDS3/ptodata/2/ina/H_COMB.seq:*
- 7: /EMC_Celerra_SIDS3/ptodata/2/ina/PCpus_COMB.seq:*
- 8: /EMC_Celerra_SIDS3/ptodata/2/ina/PP_COMB.seq:*
- 9: /EMC_Celerra_SIDS3/ptodata/2/ina/RE_COMB.seq:*
- 10: /EMC_Celerra_SIDS3/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1265	96.7	1323	US-09-259-214-1	Sequence 1, Appli
2	1265	96.7	1323	US-09-318-528-1	Sequence 1, Appli
3	1265	96.7	1323	US-09-291-931-1	Sequence 1, Appli
4	1265	96.7	1323	US-09-580-515-1	Sequence 1, Appli
5	1265	96.7	1323	US-09-866-379-1	Sequence 1, Appli
6	1263.8	96.6	1301	US-09-866-379-6	Sequence 6, Appli
7	1263.8	96.6	1301	US-09-866-379-7	Sequence 7, Appli
8	1262.2	96.5	1301	US-09-866-379-5	Sequence 5, Appli
9	1262.2	96.5	1301	US-09-866-379-9	Sequence 9, Appli
10	1233.6	94.3	1486	US-09-710-477-2	Sequence 2, Appli
11	1228	93.9	1489	US-09-540-149A-9	Sequence 9, Appli
12	1228	93.9	1489	US-10-266-041A-9	Sequence 9, Appli
13	1222.4	93.5	1486	US-09-715-477-4	Sequence 4, Appli
14	1153	88.1	1272	US-08-910-798-1	Sequence 1, Appli
15	66.2	5.1	1266	US-09-489-039A-341	Sequence 341, App
16	36.8	3.0	505	US-09-621-976-15639	Sequence 15639, A
17	36.6	2.8	2868	US-09-710-794-4	Sequence 4, Appli
18	35.8	2.7	1506	US-09-902-540-3515	Sequence 3515, Ap
19	35.8	2.7	16720	US-09-902-540-1168	Sequence 1168, Ap
20	34.4	2.6	4403765	US-09-103-840A-2	Sequence 2, Appli
21	34.4	2.6	4411529	US-09-103-840A-1	Sequence 1, Appli
22	34.2	2.6	4403765	US-09-103-840A-2	Sequence 2, Appli
23	34.2	2.6	4411529	US-09-103-840A-1	Sequence 1, Appli

24	34	2.6	601	3	US-09-949-016-109215	Sequence 109215,
25	34	2.6	601	3	US-09-949-016-109216	Sequence 109216,
26	34	2.6	601	3	US-09-949-016-109217	Sequence 109217,
27	34	2.6	998	2	US-07-885-089B-5	Sequence 5, Appli
28	34	2.6	154600	3	US-09-949-016-14757	Sequence 14757, A
29	33.8	2.6	399	3	US-09-621-976-8976	Sequence 8976, Ap
30	33.6	2.6	1359	2	US-09-860-200P-1	Sequence 1, Appli
31	33	2.5	2511	2	US-08-680-326-116	Sequence 116, App
32	33	2.5	35100	2	US-08-770-379-17	Sequence 17, Appl
33	33	2.5	35100	3	US-08-757-669A-17	Sequence 17, Appl
34	33	2.5	35100	3	US-09-230-371A-17	Sequence 17, Appl
35	33	2.5	69701	3	US-09-949-016-14187	Sequence 14187, A
36	33	2.5	73308	3	US-09-949-016-16326	Sequence 16326, A
37	32.6	2.5	3147	2	US-08-441-430-3	Sequence 3, Appli
38	32.6	2.5	4485	3	US-09-023-655-1286	Sequence 1286, Ap
39	32.6	2.5	4488	2	US-08-441-430-1	Sequence 1, Appli
40	32.6	2.5	4567	3	US-09-949-016-20	Sequence 20, Appl
41	32.6	2.5	5877	3	US-10-152-886-94	Sequence 94, Appl
42	32.4	2.5	97221	3	US-09-949-016-12755	Sequence 12755, A
43	32.2	2.5	2430	3	US-09-489-039A-3732	Sequence 3732, Ap
44	32.2	2.5	3076	3	US-09-710-794-1	Sequence 1, Appli
45	32.2	2.5	12257	3	US-09-949-016-16384	Sequence 16384, A

ALIGNMENTS

RESULT 1
US-09-259-214-1
Sequence 1, Application US/09259214A
Patent No. 6110719
GENERAL INFORMATION:
APPLICANT: Kretz, Keith
TITLE OF INVENTION: NOVEL PHYTASE
FILE REFERENCE: DIVER370-1
CURRENT APPLICATION NUMBER: US/09/259,214A
CURRENT FILING DATE: 1999-03-01
EARLIER APPLICATION NUMBER: 08/910,798
EARLIER FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1323
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: CDS
LOCATION: (1) ... (1320)
NAME/KEY: misc_feature
LOCATION: (1) ... (1323)
OTHER INFORMATION: n = A,T,C or G
US-09-259-214-1

Query Match 96.7% Score 1265; DB 3; Length 1323;
Best Local Similarity 98.0%; Pred. No. 0; Mismatches 26; Indels 0; Gaps 0;
Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY	1	ATGAAAGGATCTTAATCCATTTTATCTCTTGATTCGTTAACCCGCAATCTCA	60
DB	1	ATGAAAGGATCTTAATCCATTTTATCTCTTGATTCGTTAACCCGCAATCTCA	60
QY	61	TTTCGTGAGTGAAGCCGAGCTGAAGCTGGAAGTGTGATTTGATGATCGTATGCT	120
DB	61	TTTCGTGAGTGAAGCCGAGCTGAAGCTGGAAGTGTGATTTGATGATCGTATGCT	120
QY	121	GTGCGTGTCCAAACCAAGCCAGCACTGATGACGATGACCCGAGCATGGCA	180
DB	121	GTGCGTGTCCAAACCAAGCCAGCACTGATGACGATGACCCGAGCATGGCA	180
QY	181	ACTGCGCGGTAAACTGGGTGAAGTGAACACCGCGGTGTGAGTATGCGCTATCTC	240
DB	181	ACTGCGCGGTAAACTGGGTGAAGTGAACACCGCGGTGTGAGTATGCGCTATCTC	240

```
QY 241 GGAATTACTGCGCTCAGCGTCTGGTAGCCGACGGAATTGCTGCTTAATGTCGCTGCCG 300
Db 241 GGAATTACTCACAACGCGACGCTCTGGTAGCCGACGGAATTGCTGCGCAAAAAGGCGTCCCG 300
QY 301 CAGTCTGGTCAAGTCTGCGATTATGCTGATGTCGACGAGCGTACCCGTAACACAGCGCA 360
Db 301 CAGTCTGGTCAAGTCTGCGATTATGCTGATGTCGACGAGCGTACCCGTAACACAGCGCA 360
QY 361 GCTTCGCGCGCGGCTGCGACCTGACCTGCAATTAACCGTAACATACCAGACGATACG 420
Db 361 GCTTCGCGCGCGGCTGCGACCTGACCTGCAATTAACCGTAACATACCAGACGATACG 420
QY 421 TCAGTCCCGATCCGTTATTTATCTCTAAATACTGGCGTTGGCACTGATTAACGCG 480
Db 421 TCAGTCCCGATCCGTTATTTATCTCTAAATACTGGCGTTGGCACTGATTAACGCG 480
QY 481 AACGTACTGACGCGATCTTGAAGGCGAGAGGGTCAATTGCTGACTTTAACCGGCGAT 540
Db 481 AACGTACTGACGCGATCTTGAAGGCGAGAGGGTCAATTGCTGACTTTAACCGGCGAT 540
QY 541 TATGAAACGGGGTTTCGGAACCTGGAACGGGCTTAATTTCCGCAATCAACTGTGCG 600
Db 541 CCGGAAACGGGGTTTCGGAACCTGGAACGGGCTTAATTTCCGCAATCAACTGTGCG 600
QY 601 CTTAAACGTGAGAAACAGAGCAAAAGCTGTTCAATTACGACGACATTAACATCGGAATC 660
Db 601 CTTAAACGTGAGAAACAGAGCAAAAGCTGTTCAATTACGACGACATTAACATCGGAATC 660
QY 661 AAGGTAGCGCGGACTGTGTCTATTAAACCGGTGCGTAAAGCTTCGCAATGCTGACG 720
Db 661 AAGGTAGCGCGGACTGTGTCTATTAAACCGGTGCGTAAAGCTTCGCAATGCTGACG 720
QY 721 GAGATATTTCTCTGCAACAGCAAGGGAATGCGGAGCGGCGGTGGGAAAGATGACG 780
Db 721 GAGATATTTCTCTGCAACAGCAAGGGAATGCGGAGCGGCGGTGGGAAAGATGACG 780
QY 781 GATTCAACAGATGAAACCTTGTAAGTTGCAATACGCGCAATTTGATTGCTACAA 840
Db 781 GATTCAACAGATGAAACCTTGTAAGTTGCAATACGCGCAATTTGATTGCTACAA 840
QY 841 CGCAGCGCAGAGGTTGCGCGACCGCGCACCCCGTTATTAATGATTGATCAACAGCG 900
Db 841 CGCAGCGCAGAGGTTGCGCGACCGCGCACCCCGTTATTAATGATTGATCAACAGCG 900
QY 901 TTGAGCGCCCATTCACCGCAAAACAGGCGTATGCTGATTAACCACTTCACTGCTG 960
Db 901 TTGAGCGCCCATTCACCGCAAAACAGGCGTATGCTGATTAACCACTTCACTGCTG 960
QY 961 TTTATCGCCGACAGATATCTATCTGCAATCTCGCGCGCACTGAGCTCAACTG 1020
Db 961 TTTATCGCCGACAGATATCTATCTGCAATCTCGCGCGCACTGAGCTCAACTG 1020
QY 1021 ACGCTTCCCGGTCAAGCGGATTAACGCGCGGAGGTGTAATCTGTTGAAACGCTG 1080
Db 1021 ACGCTTCCCGGTCAAGCGGATTAACGCGCGGAGGTGTAATCTGTTGAAACGCTG 1080
QY 1081 CGTGGCGTAAGCGATTAACGCGAGTGAATGAGTTGCTGCTTCAACCTTAAG 1140
Db 1081 CGTGGCGTAAGCGATTAACGCGAGTGAATGAGTTGCTGCTTCAACCTTAAG 1140
QY 1141 CAGATGCTGATTAACCGCGCTGTCTTAATAACGCGCGCGAGAGTGAATCTGAC 1200
Db 1141 CAGATGCTGATTAACCGCGCTGTCTTAATAACGCGCGCGAGAGTGAATCTGAC 1200
QY 1201 CTGGCAGATGTAAGAGCGAAATGCGAGGCGATGTGTTGTTGAGGTTTACGCA 1260
Db 1201 CTGGCAGATGTAAGAGCGAAATGCGAGGCGATGTGTTGTTGAGGTTTACGCA 1260
QY 1261 ATCGGAATGAAGCAGATACCGCGCGTGAAGTTGAGATCTCAATC 1306
Db 1261 ATCGGAATGAAGCAGATACCGCGCGTGAAGTTGAGATCTCAATC 1306
```

```
RESULT 2
US-09-318-528-1
; Sequence 1, Application US/09318528
; Patent No. 6183740
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYLASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/318, 528
; EARLIER FILING DATE: 1999-05-25
; EARLIER APPLICATION NUMBER: 09/291, 931
; EARLIER FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910, 798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259, 214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) .. (1320)
; NAME/KEY: misc feature
; LOCATION: (1) .. (1323)
; OTHER INFORMATION: n = A,T,C or G
US-09-318-528-1
```

```
Query Match 96.7%; Score 1265; DB 3; Length 1323;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
```

```
QY 1 ATGAAAGCATCTTAATCCATTTTATCTCTTGAATTCGCTTAACCCCGCAATCTGCA 60
Db 1 ATGAAAGCATCTTAATCCATTTTATCTCTTGAATTCGCTTAACCCCGCAATCTGCA 60
QY 61 TTGCTCAGAGTGAAGCGGAGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 120
Db 61 TTGCTCAGAGTGAAGCGGAGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 120
QY 121 GTGCGTCTCCAAACCAAGGCGACGCACTGATGAGATGCAACCCGACATGCGCA 180
Db 121 GTGCGTCTCCAAACCAAGGCGACGCACTGATGAGATGCAACCCGACATGCGCA 180
QY 181 ACTGGCCGGTAAACCTGGGTGAGCTGACACCGCGCGGTGAGCTTAATGCTATCTC 240
Db 181 ACTGGCCGGTAAACCTGGGTGAGCTGACACCGCGCGGTGAGCTTAATGCTATCTC 240
QY 241 GGAATTACTGCGCTCAGCGTCTGGTAGCCGACGGAATTGCTGCTTAATGTCGCTGCCG 300
Db 241 GGAATTACTCACAACGCGACGCTCTGGTAGCCGACGGAATTGCTGCGCAAAAAGGCGTCCCG 300
QY 301 CAGTCTGGTCAAGTCTGCGATTATGCTGATGTCGACGAGCGTACCCGTAACACAGCGCA 360
Db 301 CAGTCTGGTCAAGTCTGCGATTATGCTGATGTCGACGAGCGTACCCGTAACACAGCGCA 360
QY 361 GCTTCGCGCGCGGCTGCGACCTGACCTGCAATTAACCGTAACATACCAGACGATACG 420
Db 361 GCTTCGCGCGCGGCTGCGACCTGACCTGCAATTAACCGTAACATACCAGACGATACG 420
QY 421 TCAGTCCCGATCCGTTATTTATCTCTAAATACTGGCGTTGGCACTGATTAACGCG 480
Db 421 TCAGTCCCGATCCGTTATTTATCTCTAAATACTGGCGTTGGCACTGATTAACGCG 480
QY 481 AACGTACTGACGCGATCTTGAAGGCGAGAGGGTCAATTGCTGACTTTAACCGGCGAT 540
Db 481 AACGTACTGACGCGATCTTGAAGGCGAGAGGGTCAATTGCTGACTTTAACCGGCGAT 540
QY 541 TATGAAACGGGGTTTCGGAACCTGGAACGGGCTTAATTTCCGCAATCAACTGTGCG 600
Db 541 CCGGAAACGGGGTTTCGGAACCTGGAACGGGCTTAATTTCCGCAATCAACTGTGCG 600
```

QY 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTGATTAAACGAGCATTAACATCGGAATC 660
Db 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTGATTAAACGAGCATTAACATCGGAATC 660
QY 661 AAGGTGAGCGCCGAGCTGTGTCTCATTTAACGGTGCGGTAAAGCTCGCATCAATGCTGAGC 720
Db 661 AAGGTGAGCGCCGAGCTGTGTCTCATTTAACGGTGCGGTAAAGCTCGCATCAATGCTGAGC 720
QY 721 GAGATATTTCTCTCTGCAACAGCAACAGGAAATGCCGAGCCGGGGTGGGAAAGATCAAC 780
Db 721 GAGATATTTCTCTCTGCAACAGCAACAGGAAATGCCGAGCCGGGGTGGGAAAGATCAAC 780
QY 781 GATTACACCAATGAGAAACCTTGTCTAATTGCTATTAAGCGCAATTTGATTTGCTACAA 840
Db 781 GATTACACCAATGAGAAACCTTGTCTAATTGCTATTAAGCGCAATTTGATTTGCTACAA 840
QY 841 CGCAGCGCCAGAGGTGGCCCGCAGCGCGCCACCCCGTTATTGATTGATCAAGACAGCG 900
Db 841 CGCAGCGCCAGAGGTGGCCCGCAGCGCGCCACCCCGTTATTGATTGATCAAGACAGCG 900
QY 901 TTGACGCCCCCATCAACCGCAAAACAGGCGTATGCTGATCAATTACCACTTCAAGTCTG 960
Db 901 TTGACGCCCCCATCAACCGCAAAACAGGCGTATGCTGATCAATTACCACTTCAAGTCTG 960
QY 961 TTTATCGCCGAGCAAGATCTAATCTGAGCAATCTGCGCGGCACTGAGCTCAACTG 1020
Db 961 TTTATCGCCGAGCAAGATCTAATCTGAGCAATCTGCGCGGCACTGAGCTCAACTG 1020
QY 1021 ACGCTTCCCGGTGACCGCGATTAACAGCGCGCAGAGGTGTAATCTGATTTGAAAGCTG 1080
Db 1021 ACGCTTCCCGGTGACCGCGATTAACAGCGCGCAGAGGTGTAATCTGATTTGAAAGCTG 1080
QY 1081 CGTGGCGTAAAGGATTAACAGCGAGTGAATTCAGTTTCTGCTGCTTCCAGACTTAAAG 1140
Db 1081 CGTGGCGTAAAGGATTAACAGCGAGTGAATTCAGTTTCTGCTGCTTCCAGACTTAAAG 1140
QY 1141 CAGATCGGTGATTAACCGCGCTGTCTAATTAAGCGCGCGCGAGAGGTGAATCTGAC 1200
Db 1141 CAGATCGGTGATTAACCGCGCTGTCTAATTAAGCGCGCGCGAGAGGTGAATCTGAC 1200
QY 1201 CTGGCAGGATGTAAGAGCGAAATGCGAGGCGATGTGTTGTTGCGAGTTTAAACGCA 1260
Db 1201 CTGGCAGGATGTAAGAGCGAAATGCGAGGCGATGTGTTGTTGCGAGTTTAAACGCA 1260
QY 1261 ATCGTGAATGAAGACGATACCGCGCTGTCAGTTTGAAGATCTCAATC 1306
Db 1261 ATCGTGAATGAAGACGATACCGCGCTGTCAGTTTGAAGATCTCAATC 1306

RESULT 3
US-09-291-931-1
; Sequence 1, Application US/09291931A
; Patent No. 6190897
; GENERAL INFORMATION:
; APPLICANT: Kretz, Kelch
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/291,931A
; EARLIER FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)

NAME/KEY: misc_feature
; LOCATION: (1)...(1323)
; OTHER INFORMATION: n = A,T,C or G
US-09-291-931-1

Query Match 96.7%; Score 1265; DB 3; Length 1323;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 ATGAAGGAGATCTTAATCCATTTTATCTCTTGTGATTCGTTAAACCCCGAATCTGCA 60
Db 1 ATGAAGGAGATCTTAATCCATTTTATCTCTTGTGATTCGTTAAACCCCGAATCTGCA 60
QY 61 TTGCGTCAAGGTGAGCCCGAGCTGAAGCTGGAAGTGTGTGATTTGCTAGTGTCTAGT 120
Db 61 TTGCGTCAAGGTGAGCCCGAGCTGAAGCTGGAAGTGTGTGATTTGCTAGTGTCTAGT 120
QY 121 GTGCGGTGCCCAACCAAGGCCACGCAATGATGAGAGATGTCACCCGAGCGATGGCCA 180
Db 121 GTGCGGTGCCCAACCAAGGCCACGCAATGATGAGAGATGTCACCCGAGCGATGGCCA 180
QY 181 ACCTGCGCGGTAAACTGCGGTGAGCTGACACCGCGCGGTGATGATCAATCCCTATCTC 240
Db 181 ACCTGCGCGGTAAACTGCGGTGAGCTGACACCGCGCGGTGATGATCAATCCCTATCTC 240
QY 241 GGCATTAATGCGCGGTGAGCGGTGAGCGGATGCTGCTGATTAATGTTGCTGCGCG 300
Db 241 GGCATTAATGCGCGGTGAGCGGTGAGCGGATGCTGCTGATTAATGTTGCTGCGCG 300
QY 301 CAGTGTGATGAGGTGAGATTAATGCTGATGTCAGAGCGGATACCGGTAAACAGGCCGA 360
Db 301 CAGTGTGATGAGGTGAGATTAATGCTGATGTCAGAGCGGATACCGGTAAACAGGCCGA 360
QY 361 GCTTGGCGCGCGGTGAGCTGACCTGATCTGATTAATGCTGATTAATGCTGATTAATG 420
Db 361 GCTTGGCGCGCGGTGAGCTGACCTGATCTGATTAATGCTGATTAATGCTGATTAATG 420
QY 421 TCAGTCCCGATCGCTTAATTTAATCTCTTAATACTGCGCTTTCGCAATGAGTAACGG 480
Db 421 TCAGTCCCGATCGCTTAATTTAATCTCTTAATACTGCGCTTTCGCAATGAGTAACGG 480
QY 481 AACGTGATGAGCGGATCTCTGAGAGGCGCAGAGGAGGTCAATGCTGATTAACCGGCA 540
Db 481 AACGTGATGAGCGGATCTCTGAGAGGCGCAGAGGAGGTCAATGCTGATTAACCGGCA 540
QY 541 TATCAAAAGGCGTTTCGGAAGTGAAGCGGCTTAATTTTCGCAATCAACTTGTGC 600
Db 541 TATCAAAAGGCGTTTCGGAAGTGAAGCGGCTTAATTTTCGCAATCAACTTGTGC 600
QY 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAAACGAGCATTAACATCGGAATC 660
Db 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAAACGAGCATTAACATCGGAATC 660
QY 661 AAGGTGAGCGCCGAGCTGTGTCTCATTTAACGGTGCGGTAAAGCTCGCATCAATGCTGAGC 720
Db 661 AAGGTGAGCGCCGAGCTGTGTCTCATTTAACGGTGCGGTAAAGCTCGCATCAATGCTGAGC 720
QY 721 GAGATATTTCTCTCTGCAACAGCAACAGGAAATGCCGAGCCGGGGTGGGAAAGATCAAC 780
Db 721 GAGATATTTCTCTCTGCAACAGCAACAGGAAATGCCGAGCCGGGGTGGGAAAGATCAAC 780
QY 781 GATTACACCAATGAGAAACCTTGTCTAATTGCTATTAAGCGCAATTTGATTTGCTACAA 840
Db 781 GATTACACCAATGAGAAACCTTGTCTAATTGCTATTAAGCGCAATTTGATTTGCTACAA 840
QY 841 CGCAGCGCCAGAGGTGGCCCGCAGCGCGCCACCCCGTTATTGATTGATCAAGACAGCG 900
Db 841 CGCAGCGCCAGAGGTGGCCCGCAGCGCGCCACCCCGTTATTGATTGATCAAGACAGCG 900
QY 901 TTGACGCCCCCATCAACCGCAAAACAGGCGTATGCTGATCAATTACCACTTCAAGTCTG 960
Db 901 TTGACGCCCCCATCAACCGCAAAACAGGCGTATGCTGATCAATTACCACTTCAAGTCTG 960

QY 961 TTTATCGCCGAGACAGATCTAATCTGCAAAATCTCGCGGCGGCACTGAGCTCAACTGG 1020
|
|
|
Db 961 TTTATTCGCGGACACAGATCTAATCTGCAAAATCTCGCGGCGGCACTGAGCTCAACTGG 1020
|
|
|
QY 1021 ACGCTTCCCGGTCAAGCCGGATTAACAAGCCGCGCAAGTGTGAACCTGGTTTGAACGCTGG 1080
|
|
|
Db 1021 ACGCTTCCCGGTCAAGCCGGATTAACAAGCCGCGCAAGTGTGAACCTGGTTTGAACGCTGG 1080
|
|
|
QY 1081 CGTGGCGTAAGCGATTAACAGCAGTGAATTCAGGTTTCGTGCTTCCAGACTTTACAG 1140
|
|
|
Db 1081 CGTGGCGTAAGCGATTAACAGCAGTGAATTCAGGTTTCGTGCTTCCAGACTTTACAG 1140
|
|
|
QY 1141 CAGATGCGTGAATAAAACGCGCTGTCTAATAATACGCGCCGAGAGAGTGAACCTGACC 1200
|
|
|
Db 1141 CAGATGCGTGAATAAAACGCGCTGTCTAATAATACGCGCCGAGAGAGTGAACCTGACC 1200
|
|
|
QY 1201 CTGGCAGAGATGTGAAGACCGAAATGCGCAGGCGCATGTGTGTTGGCAGGTTTAAACGCA 1260
|
|
|
Db 1201 CTGGCAGAGATGTGAAGACCGAAATGCGCAGGCGCATGTGTGTTGGCAGGTTTAAACGCA 1260
|
|
|
QY 1261 ATCGGATGAAGACCGCATACCGCGCTGCACTTGAATCTCATC 1306
|
|
|
Db 1261 ATCGGATGAAGACCGCATACCGCGCTGCACTTGAATCTCATC 1306
|
|
|

RESULT 4

US-09-580-515-1
; Sequence 1, Application US/09580515
; Patent No. 6720014
; GENERAL INFORMATION:
; APPLICANT: Kretz, Kelch
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/580, 515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1320)
; NAME/KEY: misc_feature
; LOCATION: (1)..(1323)
; OTHER INFORMATION: n = A,T,C or G
US-09-580-515-1

Query Match 96.7%; Score 1265; DB 3; Length 1323;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 ATGAAAGGATCTTAATCCATTTTATCTCTTGATTCGGTAAACCCCGCAATCTGCA 60
|
|
|
Db 1 ATGAAAGGATCTTAATCCATTTTATCTCTTGATTCGGTAAACCCCGCAATCTGCA 60
|
|
|
QY 61 TTGCGTCAAGTGAAGCCGAGCTGAAGCTGAAGTGTGATTTGTCAGTCTCATGGT 120
|
|
|
Db 61 TTGCGTCAAGTGAAGCCGAGCTGAAGCTGAAGTGTGATTTGTCAGTCTCATGGT 120
|
|
|
QY 121 GTGCGTGTCCAAACCAAGGCCAAGCACTGATGCGAGATGTCACCCCGAAGCATGGCCA 180
|
|
|
Db 121 GTGCGTGTCCAAACCAAGGCCAAGCACTGATGCGAGATGTCACCCCGAAGCATGGCCA 180
|
|
|

QY 181 ACTGCGCGGTAATAAAGTGGTAGACTGACACCGCGCGGTGGTAGTAATGCTATCTC 240
|
|
|
Db 181 ACTGCGCGGTAATAAAGTGGTAGACTGACACCGCGCGGTGGTAGTAATGCTATCTC 240
|
|
|
QY 241 GGAATTAATGAGCGTCAAGCGTCTGTAGCCGAGCGAATTTGCTCTTAATGTGCTGCCG 300
|
|
|
Db 241 GGAATTAATGAGCGTCAAGCGTCTGTAGCCGAGCGAATTTGCTCTTAATGTGCTGCCG 300
|
|
|
QY 301 CAGCTGTGACAGGTGCGGATTAATTGCTGATGTGACAGAGGTACCCGTAATAACAGGGCAA 360
|
|
|
Db 301 CAGCTGTGACAGGTGCGGATTAATTGCTGATGTGACAGAGGTACCCGTAATAACAGGGCAA 360
|
|
|
QY 361 GCTTCCCGCGCGGCTGGCACTGACTGTGCAATAACCGTACATAACCGAGCAGATACG 420
|
|
|
Db 361 GCTTCCCGCGCGGCTGGCACTGACTGTGCAATAACCGTACATAACCGAGCAGATACG 420
|
|
|
QY 421 TCCAGTCCCGATCCGTTAATTTATCTCTAATAAATCGGCGTTTCCCACTGATTAACGG 480
|
|
|
Db 421 TCCAGTCCCGATCCGTTAATTTATCTCTAATAAATCGGCGTTTCCCACTGATTAACGG 480
|
|
|
QY 481 AACGTGCTGACCGGATCTTCAGAGAGGCGAGAGGGTCAATGCTGACTTAACGGGCAT 540
|
|
|
Db 481 AACGTGCTGACCGGATCTTCAGAGAGGCGAGAGGGTCAATGCTGACTTAACGGGCAT 540
|
|
|
QY 541 TATCAAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCGCAATCAACTGTGCG 600
|
|
|
Db 541 TATCAAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCGCAATCAACTGTGCG 600
|
|
|
QY 601 CTTAAACGTGAGAAACAGAACGAAAGCTGTTCATTAAACGACGATTAACATCGGAATCTC 660
|
|
|
Db 601 CTTAAACGTGAGAAACAGAACGAAAGCTGTTCATTAAACGACGATTAACATCGGAATCTC 660
|
|
|
QY 661 AAGGTGAGCGCGGCTGTGCTCTAATAACGGGTGCGGTAAGCCTCGATCAATGCTGACG 720
|
|
|
Db 661 AAGGTGAGCGCGGCTGTGCTCTAATAACGGGTGCGGTAAGCCTCGATCAATGCTGACG 720
|
|
|
QY 721 GAGATATTTCTCTGCAACAAGCAGAGGATGCGAGCGCGGGTGGGAAGGATCAC 780
|
|
|
Db 721 GAGATATTTCTCTGCAACAAGCAGAGGATGCGAGCGCGGGTGGGAAGGATCAC 780
|
|
|
QY 781 GATTCAACAACGATGGAACACCTTGCTAAGTTGCTAACAACGCGCAATTTGATTTGCTACA 840
|
|
|
Db 781 GATTCAACAACGATGGAACACCTTGCTAAGTTGCTAACAACGCGCAATTTGATTTGCTACA 840
|
|
|
QY 841 CGCAGCGCAGAGGTGTGCGCGCGCGCGCACCCCGTTAATTAAGTTGATCAACAACGCG 900
|
|
|
Db 841 CGCAGCGCAGAGGTGTGCGCGCGCGCGCACCCCGTTAATTAAGTTGATCAACAACGCG 900
|
|
|
QY 901 TTGACGCCCATTCACCGCAAAAACAGCGGTATGTGTGAATTAACCATCTTCAAGTGTG 960
|
|
|
Db 901 TTGACGCCCATTCACCGCAAAAACAGCGGTATGTGTGAATTAACCATCTTCAAGTGTG 960
|
|
|
QY 961 TTTATCGCCGAGACAGATTAATCTGTGCAAAATCTCGCGCGCGCACTGAGCTCAACTGG 1020
|
|
|
Db 961 TTTATCGCCGAGACAGATTAATCTGTGCAAAATCTCGCGCGCGCACTGAGCTCAACTGG 1020
|
|
|
QY 1021 ACGCTTCCCGGTCAAGCCGGATTAACAAGCCGCGCAAGTGTGAACCTGGTTTGAACGCTGG 1080
|
|
|
Db 1021 ACGCTTCCCGGTCAAGCCGGATTAACAAGCCGCGCAAGTGTGAACCTGGTTTGAACGCTGG 1080
|
|
|
QY 1081 CGTGGCGTAAGCGATTAACAGCAGTGAATTCAGGTTTCGTGCTTCCAGACTTTACAG 1140
|
|
|
Db 1081 CGTGGCGTAAGCGATTAACAGCAGTGAATTCAGGTTTCGTGCTTCCAGACTTTACAG 1140
|
|
|
QY 1141 CAGATGCGTGAATAAAACGCGCTGTCTAATAATACGCGCCGAGAGAGTGAACCTGACC 1200
|
|
|
Db 1141 CAGATGCGTGAATAAAACGCGCTGTCTAATAATACGCGCCGAGAGAGTGAACCTGACC 1200
|
|
|
QY 1201 CTGGCAGAGATGTGAAGACCGAAATGCGCAGGCGCATGTGTGTTGGCAGGTTTAAACGCA 1260
|
|
|
Db 1201 CTGGCAGAGATGTGAAGACCGAAATGCGCAGGCGCATGTGTGTTGGCAGGTTTAAACGCA 1260
|
|
|
QY 1261 ATCGGATGAAGACCGCATACCGCGCTGCACTTGAATCTCATC 1306
|
|
|

Db 1261 ATCGGAATGAAGCAGCATACCGGCTGCACTTGAATCTCATC 1306

RESULT 5

US-09-866-379-1

Sequence 1, Application US/09866379

Patent No. 6853365

GENERAL INFORMATION:

APPLICANT: DIVERSA CORPORATION

APPLICANT: SHORT, Jay

APPLICANT: KRETZ, Kelch

APPLICANT: GRAY, Kevin

APPLICANT: BARTON, Nelson

APPLICANT: GARRETT, James

APPLICANT: O'DONOGHUE, Eileen

TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF

FILE REFERENCE: DIVR1370-7

CURRENT APPLICATION NUMBER: US/09/866,379

PRIOR APPLICATION NUMBER: US 09/580,515

PRIOR FILING DATE: 2000-05-25

PRIOR APPLICATION NUMBER: US 09/318,528

PRIOR FILING DATE: 1999-05-25

PRIOR APPLICATION NUMBER: US 09/291,931

PRIOR FILING DATE: 1999-04-13

PRIOR APPLICATION NUMBER: US 09/259,214

PRIOR FILING DATE: 1999-03-01

PRIOR APPLICATION NUMBER: US 08/910,798

PRIOR FILING DATE: 1997-08-13

NUMBER OF SEQ ID NOS: 10

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 1323

TYPE: DNA

ORGANISM: Escherichia coli

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)..(1323)

NAME/KEY: CDS

LOCATION: (1)..(1323)

OTHER INFORMATION:

US-09-866-379-1

Query Match 96.7%; Score 1265; DB 3; Length 1323;

Best Local Similarity 98.0%; Pred. No. 0;

Matches 1280; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 ATGAAAGCATTTTATCCCATTTTATCTTTGATTCGGTAAACCCCGCAATCTGCA 60
 Db 1 ATGAAAGCATTTTATCCCATTTTATCTTTGATTCGGTAAACCCCGCAATCTGCA 60
 QY 61 TTGCTCAGAGTGAAGCCGAGCTGAAGCTGAAGTGTGTGATGTGATGTGATGTGAT 120
 Db 61 TTGCTCAGAGTGAAGCCGAGCTGAAGCTGAAGTGTGTGATGTGATGTGATGTGAT 120
 QY 121 GTGCGTGTCCCAACCAAGGCCACGCACTGATGACAGATGTGACCCCGACGATGGCA 180
 Db 121 GTGCGTGTCCCAACCAAGGCCACGCACTGATGACAGATGTGACCCCGACGATGGCA 180
 QY 121 GTGCGTGTCCCAACCAAGGCCACGCACTGATGACAGATGTGACCCCGACGATGGCA 180
 Db 121 GTGCGTGTCCCAACCAAGGCCACGCACTGATGACAGATGTGACCCCGACGATGGCA 180
 QY 181 ACCTGGCCGGTAAACTGGGTGAGTGAACACCGCGCGGTGTGAGCTTAATCGCTATCTC 240
 Db 181 ACCTGGCCGGTAAACTGGGTGAGTGAACACCGCGCGGTGTGAGCTTAATCGCTATCTC 240
 QY 241 GGAACATTAATGAGCGTCAAGGCTGTGAGCGGACGATGCTGCTTAATGTGGCTGCGCG 300
 Db 241 GGAACATTAATGAGCGTCAAGGCTGTGAGCGGACGATGCTGCTTAATGTGGCTGCGCG 300
 QY 301 CAGTCTGATCAGAGTGCAGATTAATGCTGATGTGACGAGCGTAAACCCGTAACAGGCGAA 360
 Db 301 CAGTCTGATCAGAGTGCAGATTAATGCTGATGTGACGAGCGTAAACCCGTAACAGGCGAA 360

QY 361 GCCTTCGCGCGGCGGCTGCAACCTGATCTGTGATTAACCGTACATACCGAGGCAATAG 420
 Db 361 GCCTTCGCGCGGCGGCTGCAACCTGATCTGTGATTAACCGTACATACCGAGGCAATAG 420
 QY 421 TCCAGTCCCGATCCGTTATTTTATCTCTTAATAAATGCGGTTTGCAATGTAACGCG 480
 Db 421 TCCAGTCCCGATCCGTTATTTTATCTCTTAATAAATGCGGTTTGCAATGTAACGCG 480
 QY 481 AACGTGATGACGCAATCTCGAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
 Db 481 AACGTGATGACGCAATCTCGAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
 QY 541 TATCAACGCGGCTTTCGCAACTGGAACGGGCTTTAATTTTCCGCAATCAACTTGTGC 600
 Db 541 TATCAACGCGGCTTTCGCAACTGGAACGGGCTTTAATTTTCCGCAATCAACTTGTGC 600
 QY 601 CTTAAACGTGAAGAACGAGACGAAAGCTGTTCAATTAACGAGGCAATACATCGGAATC 660
 Db 601 CTTAAACGTGAAGAACGAGACGAAAGCTGTTCAATTAACGAGGCAATACATCGGAATC 660
 QY 661 AAGGTGAGCGCGGAGTGTCTCATTTAACCGGTGCGTAAAGCTGCAATGCTGAGG 720
 Db 661 AAGGTGAGCGCGGAGTGTCTCATTTAACCGGTGCGTAAAGCTGCAATGCTGAGG 720
 QY 721 GAGATATTTCTCTGCAACAGCAACAGGAATGCCGAGCCGCGGTGGGAAAGATCAAC 780
 Db 721 GAGATATTTCTCTGCAACAGCAACAGGAATGCCGAGCCGCGGTGGGAAAGATCAAC 780
 QY 781 GATTCAACGAGTGAACACCTTGCTAAGTTGCAATACGCGCAATTTGATTTGCTACAA 840
 Db 781 GATTCAACGAGTGAACACCTTGCTAAGTTGCAATACGCGCAATTTGATTTGCTACAA 840
 QY 841 CGCAGCGCAAGAGTTCGCGGAGCGCGCAACCCCGTATTAGATTGATCAAGACAGG 900
 Db 841 CGCAGCGCAAGAGTTCGCGGAGCGCGCAACCCCGTATTAGATTGATCAAGACAGG 900
 QY 901 TTGAGCGCCCATCCACCGCAAAACAGGCGTATGTGTGACATTAACCTTCACTACG 960
 Db 901 TTGAGCGCCCATCCACCGCAAAACAGGCGTATGTGTGACATTAACCTTCACTACG 960
 QY 961 TTTATCGCGGACAGATTAATCTGCAATCTCGGCAATCTCGGCGGCACTGAGCTCAACTGG 1020
 Db 961 TTTATCGCGGACAGATTAATCTGCAATCTCGGCAATCTCGGCGGCACTGAGCTCAACTGG 1020
 QY 1021 ACGCTTCCCGGTACGCGGATTAACAGCGCGCGAGGTGTGAACTGTGTTTGAACGCTGG 1080
 Db 1021 ACGCTTCCCGGTACGCGGATTAACAGCGCGCGAGGTGTGAACTGTGTTTGAACGCTGG 1080
 QY 1081 CGTGGCTTAAGCATTAACAGCGAGTGAATTCAGGTTTGGCTGTTTCCAGCTTTACAG 1140
 Db 1081 CGTGGCTTAAGCATTAACAGCGAGTGAATTCAGGTTTGGCTGTTTCCAGCTTTACAG 1140
 QY 1141 CAGATGCGTGAATTAAGCGCGCTGTCATTAATACCGCGCGGAGGTGAATCGAGC 1200
 Db 1141 CAGATGCGTGAATTAAGCGCGCTGTCATTAATACCGCGCGGAGGTGAATCGAGC 1200
 QY 1201 CTGCGAGATGTGAAGACGAAATGCGAGGAGCATGTGTTGTTGCGAGTTTACGCA 1260
 Db 1201 CTGCGAGATGTGAAGACGAAATGCGAGGAGCATGTGTTGTTGCGAGTTTACGCA 1260
 QY 1261 ATCGGAATGAAGCAGCATACCGGCTGCACTTGAATCTCATC 1306
 Db 1261 ATCGGAATGAAGCAGCATACCGGCTGCACTTGAATCTCATC 1306

RESULT 6

US-09-866-379-6

Sequence 6, Application US/09866379

Patent No. 6853365

GENERAL INFORMATION:

APPLICANT: DIVERSA CORPORATION

APPLICANT: SHORT, Jay

APPLICANT: KRETZ, Kelch

APPLICANT: GRAY, Kevin
APPLICANT: BARTON, Nelson
APPLICANT: GARRETT, James
APPLICANT: O'DONOGHUE, Bileen
TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
FILE REFERENCE: DIVER1370-7
CURRENT APPLICATION NUMBER: US/09/866,379
PRIOR FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: US 09/580,515
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 09/318,528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: US 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 09/259,214
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: US 08/910,798
PRIOR FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 1901
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(1901)
OTHER INFORMATION: n is any nucleotide
US-09-866-379-6

Query Match 96.6%; Score 1263.8; DB 3; Length 1901;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 1277; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1 ATGAAAGGATCTTATCCATTTTATCTTCTGATTCGGTTAACCCCGCATCTGCA 60
DB 188 ATGAAAGGATCTTATCCATTTTATCTTCTGATTCGGTTAACCCCGCATCTGCA 247
QY 61 TTGCGTCAGATGAGCCGAGCTGAAAGTGATGATGATGATGATGATGATGATGATGAT 120
DB 248 TTGCGTCAGATGAGCCGAGCTGAAAGTGATGATGATGATGATGATGATGATGATGAT 307
QY 121 GTGCGTGTCTCAACCAAGCCGACGCACTGATGATGATGATGATGATGATGATGATGATGAT 180
DB 308 GTGCGTGTCTCAACCAAGCCGACGCACTGATGATGATGATGATGATGATGATGATGATGAT 367
QY 181 ACCTGGCCGCTTAAACTGGGTGAGCTGACACCGCCGGTGGTGAATATGCGCTATCTC 240
DB 368 ACCTGGCCGCTTAAACTGGGTGAGCTGACACCGCCGGTGGTGAATATGCGCTATCTC 427
QY 241 GAGCATTACTGAGCGTCTGGTAGCCGAGATTGCTGCTTAATGTTGGCTGCGCG 300
DB 428 GAGCATTACTGAGCGTCTGGTAGCCGAGATTGCTGCTGCGCAAAAAGGGCTGCGCG 487
QY 301 CAGTCTGGTCAAGTCTGAGTATTTGCTGATGTCAGAGCTAACCCGTTAAACAAGCGCAA 360
DB 488 CAGTCTGGTCAAGTCTGAGTATTTGCTGATGTCAGAGCTAACCCGTTAAACAAGCGCAA 547
QY 361 GCTTTCCGCGCCGGCTGGCACTGATCTGTGCAATACCGTACATACCCAGCGAGATACG 420
DB 548 GCTTTCCGCGCCGGCTGGCACTGATCTGTGCAATACCGTACATACCCAGCGAGATACG 607
QY 421 TCCAGTCCCGATCGTTATTTAATCTCTTAAAACTGGCGTTTGCACTGATTAACGCG 480
DB 608 TCCAGTCCCGATCGTTATTTAATCTCTTAAAACTGGCGTTTGCACTGATTAACGCG 667
QY 481 AACGTGACTGACGCGATCTCGAGAGGCGCAGAGGGTCAATTGCTGATTAACGCGGAT 540
DB 668 AACGTGACTGACGCGATCTCGAGAGGCGCAGAGGGTCAATTGCTGATTAACGCGGAT 727
QY 541 TATCAACGCGGTTTCCGAACTGGAACGCGGTGCTTAATTTTCCGAACTCAACTTGTGC 600
DB 728 CCGCAACGCGGTTTCCGAACTGGAACGCGGTGCTTAATTTTCCGCAATCAAACTTGTGC 787

QY 601 CTTAAACGTGAAACAGAGCAAGAGCTGTTCACTTAACGAGGATTAACATCGAACTC 660
DB 788 CTTAAACGTGAAACAGAGCAAGAGCTGTTCACTTAACGAGGATTAACATCGAACTC 847
QY 661 AAGGTAGCGCCGACGTGTCTCACTTAACCGGTCCGTTAGCTTCGATCAATCTGACG 720
DB 848 AAGGTAGCGCCGACCAATGTCTCACTTAACCGGTCCGTTAGCTTCGATCAATCTGACG 907
QY 721 GAGTATTTCTCTCGCAACAGCAAGGAAATGCGGAGCCGGGTGGGAAAGATCAAC 780
DB 908 GAGTATTTCTCTCGCAACAGCAAGGAAATGCGGAGCCGGGTGGGAAAGATCAAC 967
QY 781 GATTCAACACAGTGAACACCTTCTAAGTTTGATACGCGCAATTTGATTTTCTACAA 840
DB 968 GATTCAACACAGTGAACACCTTCTAAGTTTGATACGCGCAATTTTATTTGCTACAA 1027
QY 841 CGACGCGCAGAGTTGCGCCGAGCGCCGCAACCCGTTATTAGATTGATCAAGACGCG 900
DB 1028 CGACGCGCAGAGTTGCGCCGAGCGCCGCAACCCGTTATTAGATTGATCAAGACGCG 1087
QY 901 TTGACGCCCCATCCACCGCAAAAACAGCGCTATGATGATTAACCACTTCAGTGTG 960
DB 1088 TTGACGCCCCATCCACCGCAAAAACAGCGCTATGATGATTAACCACTTCAGTGTG 1147
QY 961 TTTATGCGCCGACAGATTAATCTGCAAAATCTCGCGCGCGACCTGAGCTCAACTG 1020
DB 1148 TTTATGCGCCGACAGATTAATCTGCAAAATCTCGCGCGCGACCTGAGCTCAACTG 1207
QY 1021 ACGCTTCCCGGTGACCGGATTAACAGCCCGCAGTGGTGAATGTTGTTGAACGCTGG 1080
DB 1208 ACGCTTCCCGGTGACCGGATTAACAGCCCGCAGTGGTGAATGTTGTTGAACGCTGG 1267
QY 1081 CGTGGCTTAAGCGTAAACAGCAATGATTCAGTTTCGTGCTTCGAGACTTTCACG 1140
DB 1268 CGTGGCTTAAGCGTAAACAGCAATGATTCAGTTTCGTGCTTCGAGACTTTCACG 1327
QY 1141 CAGATGCGTATTAACAGCCGCTGTCACTTAATACGCGCCGAGAGGTAACTGACC 1200
DB 1328 CAGATGCGTATTAACAGCCGCTGTCACTTAATACGCGCCGAGAGGTAACTGACC 1387
QY 1201 CTGCGAGATGTAAGAGCGAATATGCGAGGCAATGTTGTTGCAAGTTTACGCAA 1260
DB 1388 CTGCGAGATGTAAGAGCGAATATGCGAGGCAATGTTGTTGCAAGTTTACGCAA 1447
QY 1261 ATCGTAATGAAGACGATACCGGCGTGCAGTTTGAGAT 1300
DB 1448 ATCGTAATGAAGACGATACCGGCGTGCAGTTTGAGAT 1487

RESULT 7
US-09-866-379-7
Sequence 7, Application US/09866379
Patent No. 6855365
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: SHORT, Jay
APPLICANT: KRETZ, Keith
APPLICANT: GRAY, Kevin
APPLICANT: BARTON, Nelson
APPLICANT: GARRETT, James
APPLICANT: O'DONOGHUE, Bileen
TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
FILE REFERENCE: DIVER1370-7
CURRENT APPLICATION NUMBER: US/09/866,379
PRIOR FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: US 09/580,515
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 09/318,528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: US 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 09/259,214

PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: US 08/910,798
PRIOR FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7
LENGTH: 1901
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(1901)
OTHER INFORMATION: n is any nucleotide
US-09-866-379-7

Query Match 96.6%; Score 1263.8; DB 3; Length 1901;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 1277; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

1 ATGAAAGGATCTTAATCCCATTTTATCTCTTGTGATCCGTTAACCCCGCAATCTGCA 60
188 ATGAAAGGATCTTAATCCCATTTTATCTCTTGTGATCCGTTAACCCCGCAATCTGCA 247
61 TTGCTCAGAGTGAAGCCGAGCTGAAAGTGAAGTGTGATTTGTCACTGCTGATGGT 120
248 TTGCTCAGAGTGAAGCCGAGCTGAAAGTGAAGTGTGATTTGTCACTGCTGATGGT 307
121 GTGCTGTCTCAACCAAGGCAACGCAATGATGACAGATGTCACCCAGACGATGCGCA 180
308 GTGCTGTCTCAACCAAGGCAACGCAATGATGACAGATGTCACCCAGACGATGCGCA 367
181 ACCTGGCCGGTAAACTGGGTGAGTGAACACCGCGCGTGTGAGTAAATCGCTATCTC 240
368 ACCTGGCCGGTAAACTGGGTGAGTGAACACCGCGCGTGTGAGTAAATCGCTATCTC 427
241 GGACATTAATCTGGCGTCAAGCGTCTGAGTGAAGCGCAATGCTGCTTAATGTCGCGCG 300
428 GGACATTAATCAACGCGCGTCTGAGTGAAGCGCAATGCTGCGCAATGTCGCGCG 487
301 CAGTCTGTGACAGTGCAGATTAATGCTGATGTCAGACGATACCCGTTAAACAGCGCA 360
488 CAGTCTGTGACAGTGCAGATTAATGCTGATGTCAGACGATACCCGTTAAACAGCGCA 547
361 GCTTTCGCGCGCGCGTGCACCTTGAATGTCATTAACCGTACATACCGACGATAGC 420
548 GCTTTCGCGCGCGCGTGCACCTTGAATGTCATTAACCGTACATACCGACGATAGC 607
421 TCCAGTCCCGATCCGTTATTAATCTCTAAACCTGGGTGTTGCAATGATGATACCGG 480
608 TCCAGTCCCGATCCGTTATTAATCTCTAAACCTGGGTGTTGCAATGATGATACCGG 667
481 AACGTGACAGACGATCTCTGAGAGGCGAGAGGCTCAATTGCTGATCTTACCGGCGAT 540
668 AACGTGACAGACGATCTCTGAGAGGCGAGAGGCTCAATTGCTGATCTTACCGGCGAT 727
541 TATCAACGCGCGTTCGCAACTGAAACGCGGTCTTAATTTTCCGCAATCAAACTTTGTC 600
728 CCGCAACGCGCGTTCGCAACTGAAACGCGGTCTTAATTTTCCGCAATCAAACTTTGTC 787
601 CTTAAACGTGAAGAACAGACGAAAGCTTTTCAATTAACGACGATTAACATCGAATCTC 660
788 CTTAAACGTGAAGAACAGACGAAAGCTTTTCAATTAACGACGATTAACATCGAATCTC 847
661 AAGGTGAGCGCGATGCTCTCATTTAACCGGTGCGTGAAGCTGCGATCAATGCTGACG 720
848 AAGGTGAGCGCGATGCTCTCATTTAACCGGTGCGTGAAGCTGCGATCAATGCTGACG 907
721 GAGATATTTCTCTGCAACAGCAACGAGGAATGCGGAGCGCGGTGAGGAAGATCAC 780
908 GAGATATTTCTCTGCAACAGCAACGAGGAATGCGGAGCGCGGTGAGGAAGATCAC 967
781 GATTGACACCAATGGAACCTTGTCAATTTGCTAATGCGCAATTTGATTTGCTACAA 840

DB 968 GATTGACACCAATGGAACCTTGTCAATTTGCTAATGCGCAATTTTATTTGCTACAA 1027
QY 841 CGCAGCGCAGAGTGTGCGCGCAGCGCGCACCCCGTTATTAATTTGATTAAGACAGCG 900
DB 1028 CGCAGCGCAGAGTGTGCGCGCAGCGCGCACCCCGTTATTAATTTGATTAAGACAGCG 1087
QY 901 TTGAGCGCCCATCAACCGCAAAACAGGCGTATGAGTGAACCTTACCGCTCAAGTCTG 960
DB 1088 TTGAGCGCCCATCAACCGCAAAACAGGCGTATGAGTGAACCTTACCGCTCAAGTCTG 1147
QY 961 TTATTCGCGCAGACAGATTAATCTGCAATCTCGCGCGCACTGAGGCTCAACTG 1020
DB 1148 TTATTCGCGCAGACAGATTAATCTGCAATCTCGCGCGCACTGAGGCTCAACTG 1207
QY 1021 ACCGTTCCCGTCAAGCGGATTAACAGCGCGCAAGTGTGAATCTGTTTGAACGCTG 1080
DB 1208 ACCGTTCCCGTCAAGCGGATTAACAGCGCGCAAGTGTGAATCTGTTTGAACGCTG 1267
QY 1081 CGTGGCTTAAGGATTAACAGCGAGTGAATTCAGGTTTGTGCTTCCAGACTTACAG 1140
DB 1268 CGTGGCTTAAGGATTAACAGCGAGTGAATTCAGGTTTGTGCTTCCAGACTTACAG 1327
QY 1141 CAGATGCGTGAATAAAGCGCGCTGCTCAATTAATAGCGCGCGAGAGGTGAACCTGACC 1200
DB 1328 CAGATGCGTGAATAAAGCGCGCTGCTCAATTAATAGCGCGCGAGAGGTGAACCTGACC 1387
QY 1201 CTGGAGAGTGTGAAGAGGAATGCGCAGGCGATGTGTTCTGTGCAAGTTTACGCA 1260
DB 1388 CTGGAGAGTGTGAAGAGGAATGCGCAGGCGATGTGTTCTGTGCAAGTTTACGCA 1447
QY 1261 ATCGTGAATGAAGCAGCATACCGCGCTGCACTTTGAGAT 1300
DB 1448 ATCGTGAATGAAGCAGCATACCGCGCTGCACTTTGAT 1487

RESULT 8
US-09-866-379-5
Sequence 5, Application US/09866379
Patent No. 685365
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: SHORT, Jay
APPLICANT: KRETTZ, Keith
APPLICANT: GRAY, Kevin
APPLICANT: BARTON, Nelson
APPLICANT: GARRETT, James
APPLICANT: O'DONOHUE, Bileen
TITLE OF INVENTION: RECOMBINANT BACTERIAL PHRYASES AND USES THEREOF
FILE REFERENCE: DIVER1370-7
CURRENT APPLICATION NUMBER: US/09/866,379
PRIOR FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: US 09/580,515
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 09/318,528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: US 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 09/259,214
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: US 08/910,798
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 1901
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(1901)
OTHER INFORMATION: n is any nucleotide
US-09-866-379-5

Query Match 96.5%; Score 1262.2; DB 3; Length 1901;
 Best Local Similarity 98.2%; Pred. No. 0;
 Matches 1276; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 ATGAAAGCATCTTAATCCATTTTATCTCTTGATTCGGTTAAACCCCGCAATCTGCA 60
 DB 188 ATGAAAGCATCTTAATCCATTTTATCTCTTGATTCGGTTAAACCCCGCAATCTGCA 247
 QY 61 TTCGCTGAGTGAAGCCGAGCTGAAGCTGAAGAGTGTGATGTGTAAGTCGTCATGCT 120
 DB 248 TTCGCTGAGTGAAGCCGAGCTGAAGAGTGTGATGTGTAAGTCGTCATGCT 307
 QY 121 GTGCGTGTCCAAACCAAGCCCAACGCACTGATGCAAGTGTCAACCCGAGAGCATGSCCA 180
 DB 308 GTGCGTGTCCAAACCAAGCCCAACGCACTGATGCAAGTGTCAACCCGAGAGCATGSCCA 367
 QY 181 ACCTGGCCGGTAAACTGGGTGAGCTGAACCGCCGGTGGTGAAGCTAATGCGCTATCTC 240
 DB 368 ACCTGGCCGGTAAACTGGGTGAGCTGAACCGCCGGTGGTGAAGCTAATGCGCTATCTC 427
 QY 241 GGACATTACTGGCGCTCAGGCTGTGGTAGCCGACGGAATTGTCCTTAATGTGGCTGGCCG 300
 DB 428 GGACATTAACCAAGCCGAGCTGTGGTAGCCGAGATTGCTGGCGAATAAGGGCTGGCCG 487
 QY 301 CAGTCTGTGAGTGTGCGGATTAATGCTGATGTGCAAGAGCGTACCCGTTAAACAGGCGAA 360
 DB 488 CAGTCTGTGAGTGTGCGGATTAATGCTGATGTGCAAGAGCGTACCCGTTAAACAGGCGAA 547
 QY 361 GCCTTCCGCGCGGCTGGCACTGATGTGTGCAATACCGTACATACCGAGCGAGATGCG 420
 DB 548 GCCTTCCGCGCGGCTGGCACTGATGTGTGCAATACCGTACATACCGAGCGAGATGCG 607
 QY 421 TCCAGTCCGATCGGTTATTTAATCTCTAAACATGGCGTTGGCACTGGATTAAGCGG 480
 DB 608 TCCAGTCCGATCGGTTATTTAATCTCTAAACATGGCGTTGGCACTGGATTAAGCGG 667
 QY 481 AACGTGACTGACGCGATCTCGAGAGGCGAGAGGGTCAATGCTGATTAACCGGCGAT 540
 DB 668 AACGTGACTGACGCGATCTCGAGAGGCGAGAGGGTCAATGCTGATTAACCGGCGAT 727
 QY 541 TATCAACGCGGTTGCGGAACGTGAACGCGGTCTTAATTTTCGCAATCAACTTGTGC 600
 DB 728 CCGGAACGCGGTTGCGGAACGTGAACGCGGTCTTAATTTTCGCAATCAACTTGTGC 787
 QY 601 CTTAAACGTGAGAAACAGAGAAAGCTGTCAATTAAGCGGCGCATTAACCATTCGAACTC 660
 DB 788 CTTAAACGTGAGAAACAGAGAAAGCTGTCAATTAAGCGGCGCATTAACCATTCGAACTC 847
 QY 661 AAGGTGAGCGCGACTGTGTCTCAATTAACCGGTGCGTGAAGCTTCGATCAATGCTGACG 720
 DB 848 AAGGTGAGCGCGACTGTGTCTCAATTAACCGGTGCGTGAAGCTTCGATCAATGCTGACG 907
 QY 721 GAGATATTTCTCTGCAACAAAGCAAGGAAATGCCGAGCCGGGGTGGGAAAGATCAC 780
 DB 908 GAGATATTTCTCTGCAACAAAGCAAGGAAATGCCGAGCCGGGGTGGGAAAGATCAC 967
 QY 781 GATTCAACAGGTGAACACCTGTGTAAGTTGCAATTAACGCGCAATTTGTTGTAACA 840
 DB 968 GATTCAACAGGTGAACACCTGTGTAAGTTGCAATTAACGCGCAATTTGTTGTAACA 1027
 QY 841 CGCAGCGCAGAGTGTGCGGAGCGCGCAACCCCGTTATTAGATTGATCAAGACAGCG 900
 DB 1028 CGCAGCGCAGAGTGTGCGGAGCGCGCAACCCCGTTATTAGATTGATCAAGACAGCG 1087
 QY 901 TTGAGCGCCCATTCACCGGAAAAACAGCGGTATGTGTGACATTAACCATTCAGTGTG 960
 DB 1088 TTGAGCGCCCATTCACCGGAAAAACAGCGGTATGTGTGACATTAACCATTCAGTGTG 1147
 QY 961 TTTATGCGCGGACAGATTAATTAATGTGCAAACTCCGCGCGGCACTGAGAGTCAACTGG 1020
 DB 1148 TTTATGCGCGGACAGATTAATTAATGTGCAAACTCCGCGCGGCACTGAGAGTCAACTGG 1207
 QY 1021 ACGCTTCCCGGTACAGCCGGAATTAACAGCGCGCAGGTGTGTAACCTGTGTTGAACGCTGG 1080

DB 1208 ACGTTCGCCGCTGACCGGATTAACAGCGCGCAGGTGTGTAACCTGTGTTGAACGCTGG 1267
 QY 1081 CGTGGCTTAAGCATTAACAGCGCAAGTGTGATTTGCTGTGCTTCCAGACTTTACAG 1140
 DB 1268 CGTGGCTTAAGCATTAACAGCGCAAGTGTGATTTGCTGTGCTTCCAGACTTTACAG 1327
 QY 1141 CAGATGCGGTATTAACAGCGCGCTGTCAATTAATAGCGCGCCGGAGAGGTGAACCTGACC 1200
 DB 1328 CAGATGCGGTATTAACAGCGCGCTGTCAATTAATAGCGCGCCGGAGAGGTGAACCTGACC 1387
 QY 1201 CTGGCAGATGTGAAGACGGAATGCGCAGGCAATGTGTTGTCGAGGTTTACGCA 1260
 DB 1388 CTGGCAGATGTGAAGACGGAATGCGCAGGCAATGTGTTGTCGAGGTTTACGCA 1447
 QY 1261 ATCGTAATGAAGACCATTCGCGCGTGCATTTGAGAT 1300
 DB 1448 ATCGTAATGAAGACCATTCGCGCGTGCATTTGTAAT 1487

RESULT 9
 US-09-866-379-9
 ; Sequence 9, Application US/09866379
 ; Patent No. 685365
 ; GENERAL INFORMATION:
 ; APPLICANT: DIVERSA CORPORATION
 ; APPLICANT: SHORT, Jay
 ; APPLICANT: KRETZ, Keith
 ; APPLICANT: GRAY, Kevin
 ; APPLICANT: BARTON, Nelson
 ; APPLICANT: GARRETT, James
 ; APPLICANT: O'DONOGHUE, Eileen
 ; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
 ; FILE REFERENCE: DIVER1370-7
 ; CURRENT FILING DATE: US/09/866,379
 ; PRIOR APPLICATION NUMBER: US 2001-05-24
 ; PRIOR FILING DATE: 2000-05-25
 ; PRIOR APPLICATION NUMBER: US 09/318,528
 ; PRIOR FILING DATE: 1999-05-25
 ; PRIOR APPLICATION NUMBER: US 09/291,931
 ; PRIOR FILING DATE: 1999-04-13
 ; PRIOR APPLICATION NUMBER: US 09/259,214
 ; PRIOR FILING DATE: 1999-03-01
 ; PRIOR APPLICATION NUMBER: US 08/910,798
 ; PRIOR FILING DATE: 1997-08-13
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 9
 ; LENGTH: 1901
 ; TYPE: DNA
 ; ORGANISM: Escherichia coli
 ; FEATURES:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)..(1901)
 ; OTHER INFORMATION: n is any nucleotide
 US-09-866-379-9

Query Match 96.5%; Score 1262.2; DB 3; Length 1901;
 Best Local Similarity 98.2%; Pred. No. 0;
 Matches 1276; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 ATGAAAGCATCTTAATCCATTTTATCTCTTGATTCGGTTAAACCCCGCAATCTGCA 60
 DB 188 ATGAAAGCATCTTAATCCATTTTATCTCTTGATTCGGTTAAACCCCGCAATCTGCA 247
 QY 61 TTCGCTGAGTGAAGCCGAGCTGAAGCTGAAGAGTGTGATGTGTAAGTCGTCATGCT 120
 DB 248 TTCGCTGAGTGAAGCCGAGCTGAAGAGTGTGATGTGTAAGTCGTCATGCT 307
 QY 121 GTGCGTGTCCAAACCAAGCCCAACGCACTGATGCAAGTGTCAACCCGAGAGCATGSCCA 180
 DB 308 GTGCGTGTCCAAACCAAGCCCAACGCACTGATGCAAGTGTCAACCCGAGAGCATGSCCA 367

181 ACCGCGCGGTTAAACTGGGTGAGCTGACACCGCGCGGTGAGCTTAATGCGTATCTC 240
 368 ACCGCGCGGTTAAACTGGGTGAGCTGACACCGCGCGGTGAGCTTAATGCGTATCTC 427
 241 GGACATTAATGAGCGTCAAGCGTCTGATGAGCGGCGGATTCGCTTAATGAGCGTCCG 300
 428 GGACATTAATCAACGCGGATGCTGATGAGCGGCGGATTCGCTTAATGAGCGTCCG 487
 301 CAGTCTGATGAGCGGATGCTGATGAGCGGCGGATTCGCTTAATGAGCGTCCG 360
 488 CAGTCTGATGAGCGGATGCTGATGAGCGGCGGATTCGCTTAATGAGCGTCCG 547
 361 GCGTTCGCGCGGCGGATGCTGATGAGCGGCGGATTCGCTTAATGAGCGTCCG 420
 548 GCGTTCGCGCGGCGGATGCTGATGAGCGGCGGATTCGCTTAATGAGCGTCCG 607
 421 TCCAGTCCGCGGATGCTGATGAGCGGCGGATTCGCTTAATGAGCGTCCG 480
 608 TCCAGTCCGCGGATGCTGATGAGCGGCGGATTCGCTTAATGAGCGTCCG 667
 481 AACGTGATGAGCGGATGCTGATGAGCGGCGGATTCGCTTAATGAGCGTCCG 540
 668 AACGTGATGAGCGGATGCTGATGAGCGGCGGATTCGCTTAATGAGCGTCCG 727
 541 TATCAAAAGCGGTTGCGGATGAGCGGCGGATTCGCTTAATGAGCGTCCG 600
 728 CCGGCAAAAGCGGTTGCGGATGAGCGGCGGATTCGCTTAATGAGCGTCCG 787
 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGAGGCAATTAACGAGATC 660
 788 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGAGGCAATTAACGAGATC 847
 661 AAGGTGAGCGGATGCTGATGAGCGGCGGATTCGCTTAATGAGCGTCCG 720
 848 AAGGTGAGCGGATGCTGATGAGCGGCGGATTCGCTTAATGAGCGTCCG 907
 721 GAGATATTTCTCTGCAACAAAGCAAGGAAATGCGGAGCGGCGGATTCGCTTAATGAG 780
 908 GAGATATTTCTCTGCAACAAAGCAAGGAAATGCGGAGCGGCGGATTCGCTTAATGAG 967
 781 GATTCAACAGTGAACCACTTGTGATTAACGAGGCAATTAACGAGATTCGCTTAATGAG 840
 968 GATTCAACAGTGAACCACTTGTGATTAACGAGGCAATTAACGAGATTCGCTTAATGAG 1027
 841 CGCAGCGCAAGAGGTTGCGGAGCGGCGGATTCGCTTAATGAGTTCGCTTAATGAG 900
 1028 CGCAGCGCAAGAGGTTGCGGAGCGGCGGATTCGCTTAATGAGTTCGCTTAATGAG 1087
 901 TTGAGCGGCGGATTCGCTTAATGAGTTCGCTTAATGAGTTCGCTTAATGAG 960
 1088 TTGAGCGGCGGATTCGCTTAATGAGTTCGCTTAATGAGTTCGCTTAATGAG 1147
 961 TTATCGCGCGGATTCGCTTAATGAGTTCGCTTAATGAGTTCGCTTAATGAG 1020
 1148 TTATCGCGCGGATTCGCTTAATGAGTTCGCTTAATGAGTTCGCTTAATGAG 1207
 1021 ACGCTTCGCGGATTCGCTTAATGAGTTCGCTTAATGAGTTCGCTTAATGAG 1080
 1208 ACGCTTCGCGGATTCGCTTAATGAGTTCGCTTAATGAGTTCGCTTAATGAG 1267
 1081 CGTCCGCTTAAGGATTAACGAGGATTCGCTTAATGAGTTCGCTTAATGAG 1140
 1268 CGTCCGCTTAAGGATTAACGAGGATTCGCTTAATGAGTTCGCTTAATGAG 1327
 1141 CAGATCGGATTAACGAGGATTCGCTTAATGAGTTCGCTTAATGAGTTCGCTTAATGAG 1200
 1328 CAGATCGGATTAACGAGGATTCGCTTAATGAGTTCGCTTAATGAGTTCGCTTAATGAG 1387
 1201 CTGGCAGGATTAACGAGGATTCGCTTAATGAGTTCGCTTAATGAGTTCGCTTAATGAG 1260
 1388 CTGGCAGGATTAACGAGGATTCGCTTAATGAGTTCGCTTAATGAGTTCGCTTAATGAG 1447

QY 1261 ATCGTAATGAGCAAGCATACCGCGGTGAGTTCGAT 1300
 Db 1448 ATCGTAATGAGCAAGCATACCGCGGTGAGTTCGAT 1487

RESULT 10
 US-09-715-477-2
 ; Sequence 2, Application US/09715477
 ; Patent No. 6841370
 ; GENERAL INFORMATION:
 ; APPLICANT: Lei, Xingen
 ; TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYTASE
 ; FILE REFERENCE: 19603/4031
 ; CURRENT APPLICATION NUMBER: US/09/715,477
 ; CURRENT FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/166,179
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 1486
 ; TYPE: DNA
 ; ORGANISM: Escherichia coli
 US-09-715-477-2

Query Match 94.3%; Score 1233.6; DB 3; Length 1486;
 Best Local Similarity 97.0%; Pred. No. 0;
 Matches 1257; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 1 ATGAAAGGATCTTAATCCATTTTATCTGTTGATTCGTTAACCCGCAATCTGCA 60
 Db 188 ATGAAAGGATCTTAATCCATTTTATCTGTTGATTCGTTAACCCGCAATCTGCA 247
 QY 61 TTGCTCAAGTGAAGCCGGAAGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 120
 Db 248 TTGCTCAAGTGAAGCCGGAAGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 307
 QY 121 GTGCGTCTCCACCAAGCCGCAAGCTGATGAGATGTCACCCGAGAGCGATGCGCA 180
 Db 308 GTGCGTCTCCACCAAGCCGCAAGCTGATGAGATGTCACCCGAGAGCGATGCGCA 367
 QY 181 ACCTGGCGGTTAACTGGGTGAGCTGACACCGCGCGGTGAGCTTAATGCGTATCTC 240
 Db 368 ACCTGGCGGTTAACTGGGTGAGCTGACACCGCGCGGTGAGCTTAATGCGTATCTC 427
 QY 241 GGACATTAATGAGCGTCAAGCGTCTGATGAGCGGCGGATTCGCTTAATGAGCGTCCG 300
 Db 428 GGACATTAATCAACGCGGATGCTGATGAGCGGCGGATTCGCTTAATGAGCGTCCG 487
 QY 301 CAGTCTGATGAGCGGATGCTGATGAGCGGCGGATTCGCTTAATGAGCGTCCG 360
 Db 488 CAGTCTGATGAGCGGATGCTGATGAGCGGCGGATTCGCTTAATGAGCGTCCG 547
 QY 361 GCGTTCGCGCGGCGGATGCTGATGAGCGGCGGATTCGCTTAATGAGCGTCCG 420
 Db 548 GCGTTCGCGCGGCGGATGCTGATGAGCGGCGGATTCGCTTAATGAGCGTCCG 607
 QY 421 TCCAGTCCGCGGATGCTGATGAGCGGCGGATTCGCTTAATGAGCGTCCG 480
 Db 608 TCCAGTCCGCGGATGCTGATGAGCGGCGGATTCGCTTAATGAGCGTCCG 667
 QY 481 AACGTGATGAGCGGATGCTGATGAGCGGCGGATTCGCTTAATGAGCGTCCG 540
 Db 668 AACGTGATGAGCGGATGCTGATGAGCGGCGGATTCGCTTAATGAGCGTCCG 727
 QY 541 TATCAAAAGCGGTTGCGGATGAGCGGCGGATTCGCTTAATGAGCGTCCG 600
 Db 728 CCGCAAAAGCGGTTGCGGATGAGCGGCGGATTCGCTTAATGAGCGTCCG 787
 QY 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGAGGCAATTAACGAGATC 660
 Db 788 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGAGGCAATTAACGAGATC 847

QY 661 AAGGTGACGCGCATGTGTCTCATTTAACCGGTGCGGTAAGCTCGCATCAATCTGACG 720
DB 848 AAGGTGACGCGCGCATGTGTCTCATTTAACCGGTGCGGTAAGCTCGCATCAATCTGACG 907
QY 721 GAGATATTTCTCTGCAACAAGCAAGGAAATGCCGAGCCGGGGTGGGAAAGATCAAC 780
DB 908 GAAATATTTCTCTGCAACAAGCAAGGAAATGCCGAGCCGGGGTGGGAAAGATCAAC 967
QY 781 GATTCAACCAAGTGAACAACCTTGTGAATTTGCAATAACGCGCAATTTGATTTGCTCA 840
DB 968 GATTCAACCAAGTGAACAACCTTGTGAATTTGCAATAACGCGCAATTTGATTTGCTCA 1027
QY 841 CGACGCGCAAGATTTGCGCGAGCGCGCAACCCGTTATTAATTTGATTAAGACAGACG 900
DB 1028 CGACGCGCAAGATTTGCGCGAGCGCGCAACCCGTTATTAATTTGATTAAGACAGACG 1087
QY 901 TTGACGCGCATTCACCGCAAAAACAGCGATATGTTGATTAACCACTTCAGTGTG 960
DB 1088 TTGACGCGCATTCACCGCAAAAACAGCGATATGTTGATTAACCACTTCAGTGTG 1147
QY 961 TTTATGCGCGGACAGATCAATCTGCAATCTGCGGCGGCGACCTGAGACTGACG 1020
DB 1148 TTTATGCGCGGACAGATCAATCTGCAATCTGCGGCGGCGACCTGAGACTGACG 1207
QY 1021 ACGCTTCCCGGTCAAGCGGATTAACAGCGCGCAGGTGTAATCTGTGTTGAACGCTG 1080
DB 1208 ACGCTTCCCGGTCAAGCGGATTAACAGCGCGCAGGTGTAATCTGTGTTGAACGCTG 1267
QY 1081 CGTGGCTAAGCGATTAACAGCGAGTGAATTCAGTTTCTGTGTTCAAGCTTTACAG 1140
DB 1268 CGTGGCTAAGCGATTAACAGCGAGTGAATTCAGTTTCTGTGTTCAAGCTTTACAG 1327
QY 1141 CAGATGCGTGAATAAAGCGCGTGTCAATAATAGCGCGCGCGAGAGTGAACCTGACC 1200
DB 1328 CAGATGCGTGAATAAAGCGCGTGTCAATAATAGCGCGCGCGAGAGTGAACCTGACC 1387
QY 1201 CTGGCAGATGTGAAGAGCGAATGCGAGGCAATGTTCGTTGCGAGTTTACGCA 1260
DB 1388 CTGGCAGATGTGAAGAGCGAATGCGAGGCAATGTTCGTTGCGAGTTTACGCA 1447
QY 1261 ATCGTGAATGAAGCAGCATACCGGCGTGCAGTTTG 1296
DB 1448 ATCGTGAATGAAGCAGCATACCGGCGTGCAGTTTG 1483

RESULT 11
US-09-540-149A-9
; Sequence 9, Application US/09540149A
; Patent No. 6511699
; GENERAL INFORMATION:
; APPLICANT: lei. Xingen
; TITLE OF INVENTION: ENZYMES WITH IMPROVED PHYTASE ACTIVITY
; FILE REFERENCE: 19603/2791
; CURRENT APPLICATION NUMBER: US/09/540.149A
; PRIOR FILING DATE: 2000-03-31
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patemint Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1489
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-540-149A-9

Query Match 93.9%; Score 1228; DB 3; Length 1489;
Best Local Similarity 96.5%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 1 ATGAAGGATCTTAATCCATTTTATCTCTTGATTCGTTAACCCCGCAATCTGCA 60
DB 182 ATGAAGGATCTTAATCCATTTTATCTCTTGATTCGTTAACCCCGCAATCTGCA 241

QY 61 TTGCTCAGAGTAGACCGGAGCTGAAGCTGGAAGTGTGTGATTTGTCACTGTCATGCT 120
DB 242 TTGCTCAGAGTAGACCGGAGCTGAAGCTGGAAGTGTGTGATTTGTCACTGTCATGCT 301
QY 121 GTGGGTCTTCAACCAAGGCGACGCACTGATGAGATGTCAACCCAGACGATGGCCA 180
DB 302 GTGGGTCTTCAACCAAGGCGACGCACTGATGAGATGTCAACCCAGACGATGGCCA 361
QY 181 ACCGTGGCGGTTAAACCTGGGTGAGCTGACACCGCGCGGTGTGAGTAACTGCTATCTC 240
DB 362 ACCGTGGCGGTTAAACCTGGGTGAGCTGACACCGCGGTGTGAGTAACTGCTATCTC 421
QY 241 GACATTAATCTGCGGTCAAGCTGTGTAAGCGGATTTGCTGCTTAATGTGCTGCGC 300
DB 422 GACATTAATCTGCGGTCAAGCTGTGTAAGCGGATTTGCTGCTTAATGTGCTGCGC 481
QY 301 CAGTCTGTCAAGTCTGCGATTTATTTGCTGATGTGACGAGGTACCCCTAATAACAGGGAA 360
DB 482 CAGTCTGTCAAGTCTGCGATTTATTTGCTGATGTGACGAGGTACCCCTAATAACAGGGAA 541
QY 361 GCCTTGCGCGCGGGCTGGGACCTGACCTGCAATTAACGTAATACCCAGGAGATACG 420
DB 542 GCCTTGCGCGCGGGCTGGGACCTGACCTGCAATTAACGTAATACCCAGGAGATACG 601
QY 421 TCCAGTCCCGATCCGTTATTTAACTCTTAATAAATGCGCTTTGCCAATGATTAACGCG 480
DB 602 TCCAGTCCCGATCCGTTATTTAACTCTTAATAAATGCGCTTTGCCAATGATTAACGCG 661
QY 481 AACGTGACTGACCGGATCTCTGAGAGGGCAGAGGGTCAATTTGCTGACTTTACCGGGCAT 540
DB 662 AACGTGACTGACCGGATCTCTGAGAGGGCAGAGGGTCAATTTGCTGACTTTACCGGGCAT 721
QY 541 TATCAACCGCGGTTTGCGCAATGGAACGAGGTCTTAATTTTCCGCAATCAAACTTGTGC 600
DB 722 CCGCAACCGCGGTTTGCGCAATGGAACGAGGTCTTAATTTTCCGCAATCAAACTTGTGC 781
QY 601 CTTAAACGTGAAGAAACAGAACGAAAGCTGTTCATTAACGAGGATTAACATCGGAATC 660
DB 782 CTTAAACGTGAAGAAACAGAACGAAAGCTGTTCATTAACGAGGATTAACATCGGAATC 841
QY 661 AAGGTGACGCGCATGTGTCTCATTTAACCGGTGCGGTAAGCTCGCATCAATCTGACG 720
DB 842 AAGGTGACGCGCATGTGTCTCATTTAACCGGTGCGGTAAGCTCGCATCAATCTGACG 901
QY 721 GAGATATTTCTCTGCAACAAGCAAGGAAATGCCGAGCCGGGGTGGGAAAGATCAAC 780
DB 902 GAAATATTTCTCTGCAACAAGCAAGGAAATGCCGAGCCGGGGTGGGAAAGATCAAC 961
QY 781 GATTCAACCAAGTGAACAACCTTGTGAATTTGCAATAACGCGCAATTTGATTTGCTCA 840
DB 962 GATTCAACCAAGTGAACAACCTTGTGAATTTGCAATAACGCGCAATTTGATTTGCTCA 1021
QY 841 CGACGCGCAAGATTTGCGCGAGCGCGCAACCCGTTATTAATTTGATTAAGACAGACG 900
DB 1022 CGACGCGCAAGATTTGCGCGAGCGCGCAACCCGTTATTAATTTGATTAAGACAGACG 1081
QY 901 TTGACGCGCATTCACCGCAAAAACAGCGATATGTTGATTAACCACTTCAGTGTG 960
DB 1082 TTGACGCGCATTCACCGCAAAAACAGCGATATGTTGATTAACCACTTCAGTGTG 1141
QY 961 TTTATGCGCGGACAGATCAATCTGCAATCTGCGGCGGCGACCTGAGACTGACG 1020
DB 1142 TTTATGCGCGGACAGATCAATCTGCAATCTGCGGCGGCGACCTGAGACTGACG 1201
QY 1021 ACGCTTCCCGGTCAAGCGGATTAACAGCGCGCAGGTGTAATCTGTGTTGAACGCTG 1080
DB 1202 ACGCTTCCCGGTCAAGCGGATTAACAGCGCGCAGGTGTAATCTGTGTTGAACGCTG 1261
QY 1081 CGTGGCTAAGCGATTAACAGCGAGTGAATTCAGTTTCTGTGTTCAAGCTTTACAG 1140
DB 1262 CGTGGCTAAGCGATTAACAGCGAGTGAATTCAGTTTCTGTGTTCAAGCTTTACAG 1321
QY 1141 CAGATGCGTGAATAAAGCGCGTGTCAATAATAGCGCGCGGAGAGTGAACCTGACC 1200

```
Db 1322 CAGAGCGGATGAAACGCGGATCATTAATACGCGCCGCGAGAGGTGAACGAC 1381
Qy 1201 CTGGCAGAGTGTGAAGAGGAAATGCGCAGGCGATGTGTTGTTGCGAGTTTAA 1260
Db 1382 CTGGCAGAGTGTGAAGAGGAAATGCGCAGGCGATGTGTTGTTGCGAGTTTAA 1441
Qy 1261 ATCGTAATGAAGACGCGATACCGGCGTGCAGTTTGAGAT 1300
Db 1442 ATCGTAATGAAGCGCGCATACCGGCGTGCAGTTTGTAAT 1481

RESULT 12
US-10-266-041A-9
; Sequence 9, Application US/10266041A
; Patent No. 6974690
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: PHOSPHATASES WITH IMPROVED PHYTASE ACTIVITY
; FILE REFERENCE: 19603/2798
; CURRENT APPLICATION NUMBER: US/10/266,041A
; PRIOR FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: 60/127,032
; PRIOR FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 09/540,149
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1489
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-266-041A-9

Query Match 93.9%; Score 1228; DB 4; Length 1489;
Best Local Similarity 96.5%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Qy 1 ATGAAAGCATCTTAATCCCATTTTATCTCTTGATTCGTTAAACCCCGCAATCTGCA 60
Db 182 ATGAAAGCATCTTAATCCCATTTTATCTCTTGATTCGTTAAACCCCGCAATCTGCA 241
Qy 61 TTGCGTCAGAGTGAAGCGCGAGCGTGAAGCTGAAGTGTGTAATGTGTCATGTGAT 120
Db 242 TTGCGTCAGAGTGAAGCGCGAGCGTGAAGCTGAAGTGTGTAATGTGTCATGTGAT 301
Qy 121 GTGCGTGTCTCAACCAAGGCGCAAGCACTGATGACAGATGTCAACCCGCAATGCGCA 180
Db 302 GTGCGTGTCTCAACCAAGGCGCAAGCACTGATGACAGATGTCAACCCGCAATGCGCA 361
Qy 181 ACCTGGCCGGTAAACTGGGTGAGCTGACACCGCGGTGTGAGCTAATGCGCTATCTC 240
Db 362 ACCTGGCCGGTAAACTGGGTGAGCTGACACCGCGGTGTGAGCTAATGCGCTATCTC 421
Qy 241 GGAATTAATGCGGCTGAGCTGCTGAGAGCGCAAGATTGCTCTTAAATGTTGCTGCCG 300
Db 422 GGAATTAATGCGGCTGAGCTGCTGAGAGCGCAAGATTGCTCTTAAATGTTGCTGCCG 481
Qy 301 CAGTCTGTGTCAGAGTGCAGATTAATGCTGATGTGACAGAGCTAACCGGTAAACAGGCGAA 360
Db 482 CAGTCTGTGTCAGAGTGCAGATTAATGCTGATGTGACAGAGCTAACCGGTAAACAGGCGAA 541
Qy 361 GCGTTGCGCGCGGCGGTGCGACCTGACTGTGCAATPACCGTAACATACCGAGCAGATNAG 420
Db 542 GCGTTGCGCGCGGCGGTGCGACCTGACTGTGCAATPACCGTAACATACCGAGCAGATNAG 601
Qy 421 TCCAGTCCCGATCCGTTATTTAATCCTCAAAAACCTGGGTTTGCAATGGAATTAACGG 480
Db 602 TCCAGTCCCGATCCGTTATTTAATCCTCAAAAACCTGGGTTTGCAATGGAATTAACGG 661
Qy 481 AACGTGACTGACGCGATCTCTGAGAGGCGAGAGGGTCAATTGCTGACTTTTACGGGCGAT 540
Db 662 AACGTGACTGACGCGATCTCTGAGAGGCGAGAGGGTCAATTGCTGACTTTTACGGGCGAT 721
```

```
Qy 541 TATCAACGGCGTTTCGGAACTGGAACGGGTCCTAATTTTCCGCAATCAACTTGTGC 600
Db 722 CGGCAAACGGGTTTCGGAACTGGAACGGGTCCTAATTTTCCGCAATCAACTTGTGC 781
Qy 601 CTTAAACGTGAGAAACAGAGCAAGAGCTGTTCAATTAACGACGATTAACATCGAATCTC 660
Db 782 CTTAAACGTGAGAAACAGAGCAAGAGCTGTTCAATTAACGACGATTAACATCGAATCTC 841
Qy 661 AAGGTAGCGCGGACTGTGTCTCATTAACCGGTGCGTAAAGCTTCGATCAATGCTGACG 720
Db 842 AAGGTAGCGCGGACTGTGTCTCATTAACCGGTGCGTAAAGCTTCGATCAATGCTGACG 901
Qy 721 CAGATATTTCTTCCTGCAACAAAGCAACAGGAAATGCGGAGCCGGGTGGGAAAGTCAAC 780
Db 902 GAAATATTTCTTCCTGCAACAAAGCAACAGGAAATGCGGAGCCGGGTGGGAAAGTCAAC 961
Qy 781 GATTCAACAGTGAACACCTGTGCTAAGTTTGCAACGCGCAATTTGATTTTGCTACAA 840
Db 962 GATTCAACAGTGAACACCTGTGCTAAGTTTGCAACGCGCAATTTGATTTTGCTACAA 1021
Qy 841 CGCAGCGCAGAGTTGCCCGGACCGCGCAACCCGTTATTAGATTGATCAAGACAGCG 900
Db 1022 CGCAGCGCAGAGTTGCCCGGACCGCGCAACCCGTTATTGATTTGATCAAGACAGCG 1081
Qy 901 TTGAGCGCCCATCCACCGCAAAAACAGCGCTATGTGATTAATCCCACTTCAAGTCTG 960
Db 1082 TTGAGCGCCCATCCACCGCAAAAACAGCGCTATGTGATTAATCCCACTTCAAGTCTG 1141
Qy 961 TTATCGCCGGAACAGTACTAATCTGCAATATCTGCGAGGCGCACTGAGCTCAACTGG 1020
Db 1142 TTATCGCCGGAACAGTACTAATCTGCAATATCTGCGAGGCGCACTGAGCTCAACTGG 1201
Qy 1021 ACGCTTCCCGGTCAACCGGATTAACAACCGCGCAGGTGTGAATCTGTGTTGAACGCTGG 1080
Db 1202 ACGCTTCCCGGTCAACCGGATTAACAACCGCGCAGGTGTGAATCTGTGTTGAACGCTGG 1261
Qy 1081 CGTGGCTTAACGATTAACGCGCATGTGATTAAGTTTGCTGTGTTTCCCACTTTACAG 1140
Db 1262 CGTGGCTTAACGATTAACGCGCATGTGATTAAGTTTGCTGTGTTTCCCACTTTACAG 1321
Qy 1141 CAGATCGTGAATTAACCGCGCTGTCATTAATACCGCGCCGCGAGAGGTGAACCTGACC 1200
Db 1322 CAGATCGTGAATTAACCGCGCTGTCATTAATACCGCGCCGCGAGAGGTGAACCTGACC 1381
Qy 1201 CTGGCAGAGTGTGAAGAGGCAATGCGCAGGCGATGTGTTGTCAGGTTTAAACGCA 1260
Db 1382 CTGGCAGAGTGTGAAGAGGCAATGCGCAGGCGATGTGTTGTCAGGTTTAAACGCA 1441
Qy 1261 ATCGTAATGAAGACGCGATACCGGCGTGCAGTTTGAGAT 1300
Db 1442 ATCGTAATGAAGCGCGCATACCGGCGTGCAGTTTGTAAT 1481

RESULT 13
US-09-715-477-4
; Sequence 4, Application US/09715477
; Patent No. 6841370
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYTASE
; FILE REFERENCE: 19603/4031
; CURRENT APPLICATION NUMBER: US/09/715,477
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,179
; PRIOR FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1486
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-715-477-4
```

Query Match 93.5%; Score 1222.4; DB 3; Length 1486;
 Best Local Similarity 96.5%; Pred. No. 0;
 Matches 1250; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

```

QY 1 ATGAAAGCGATCTTAATCCCATTTTATCTCTTGATTCGCTTAACCCCGCAATCTGCA 60
DB 188 ATGAAAGCGATCTTAATCCCATTTTATCTCTTGATTCGCTTAACCCCGCAATCTGCA 247
QY 61 TTGCGTCAGAGTGAAGCCGAGAGCTGAAGTGTGTGATTTGTCAGTCTGTCATGCT 120
DB 248 TTGCGTCAGAGTGAAGCCGAGAGCTGAAGTGTGTGATTTGTCAGTCTGTCATGCT 307
QY 121 GTGCGTCTCCAAACCAAGCCCAAGCACTGATGCAAGATGTCACCCCAAGAGTGGCA 180
DB 308 GTGCGTCTCCAAACCAAGCCCAAGCACTGATGCAAGATGTCACCCCAAGAGTGGCA 367
QY 181 ACCTGCGCGGTAAACCTGGGTGAAGCTGACACCGCGCGGTGAGCTTAATGCGCTATCTC 240
DB 368 ACCTGCGCGGTAAACCTGGGTGAGCTGACACACGCGGTGAGCTTAATGCGCTATCTC 427
QY 241 GGAATTAATGAGCGCTGACGCTCTGTGTAGCCGAGGATTCGCTTAATGTCGTCGCCG 300
DB 428 GGAATTAATGAGCGCGCGCGCTGTGTGTAGCCGAGGATTCGCGGAAAGGCTGCCG 487
QY 301 CAGTCTGATGAGTGTGCGATTAATGCTGATGTCAGACGAGCTACCCGTTAAACAGGCGAA 360
DB 488 CAGCTGTGATGAGTGTGCGATTAATGCTGATGTCAGACGAGCTACCCGTTAAACAGGCGAA 547
QY 361 GCCTTCGCGCGCGGTGGCACTGATCTGCAATTAACCGTACATCCAGGAGATACG 420
DB 548 GCCTTCGCGCGCGGTGGCACTGATCTGCAATTAACCGTACATCCAGGAGATACG 607
QY 421 TCCAGTCCCGATCCGTTATTTAACTCTTAATAACCTGCGCTTCCCACTGGAATACGCG 480
DB 608 TCCAGTCCCGATCCGTTATTTAACTCTTAATAACCTGCGCTTCCCACTGGAATACGCG 667
QY 481 AACGTGACTGACGAGATCTCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
DB 668 AACGTGACTGACGAGATCTCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 727
QY 541 TATTAACACGCGCTTCCGGAACCTGGAACGCGGTCTTAATTTTCCGCAATCAAATTGTC 600
DB 728 CCGCAAAACGCGCTTCCGGAACCTGGAACGCGGTCTTAATTTTCCGCAATCAAATTGTC 787
QY 601 CTTAAGCTGAGAAACGAGAAAGCTGCTTAATTAACGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
DB 788 CTTAAGCTGAGAAACGAGAAAGCTGCTTAATTAACGAGGAGGAGGAGGAGGAGGAGGAGGAG 847
QY 661 AAGGTGAGCGCGGACTGTGCTCATTTAAACCGGTGCGTAACCTCGCATCATGCTGACG 720
DB 848 AAGGTGAGCGCGGACTGTGCTCATTTAAACCGGTGCGTAACCTCGCATCATGCTGACG 907
QY 721 GAGATATTTCTCTGCAACAAGCAACGAGAAATGCCGAGCGCGGAGTGGGAAAGATCAC 780
DB 908 GAAATATTTCTCTGCAACAAGCAACGAGAAATGCCGAGCGCGGAGTGGGAAAGATCAC 967
QY 781 GATTCAACACGAGTGAACCTGCTAAGTTTGATTAACGCGCAATTGATTTGCTCAAA 840
DB 968 GATTCAACACGAGTGAACCTGCTAAGTTTGATTAACGCGCAATTGATTTGCTCAAA 1027
QY 841 CGCAGCGCAGAGTTCGCGGAGCGCGCAGCCCGTATTAAGATTGATCAAGACGCG 900
DB 1028 CGCAGCGCAGAGTTCGCGGAGCGCGCAGCCCGTATTAAGATTGATCAAGACGCG 1087
QY 901 TTGACGCCCATTCACCGCAAAAACAGGCGTATGAGTGTGACATTAACCATTCAGTGTG 960
DB 1088 TTGACGCCCATTCACCGCAAAAACAGGCGTATGAGTGTGACATTAACCATTCAGTGTG 1147
QY 961 TTTATCGCGGAGCAGATACATTAATCTGGCAATCTCGCGCGCGGAGCAGCTGAGACTCAAC 1020
DB 1148 TTTATCGCGGAGCAGATACATTAATCTGGCAATCTCGCGCGCGGAGCAGCTGAGACTCAAC 1207

```

```

QY 1021 ACCGTTCCCGTCAGCCGAGTAAACAGCCGCGCAGTGTGTAAGTGTGTAACGCTGG 1080
DB 1208 ACCGTTCCAGATCAGCCGAGTAAACAGCCGCGCAGTGTGTAAGTGTGTAACGCTGG 1267
QY 1081 CGTGGCTTAAGCATTAACAGCCAGTGTGATTCAGTTCGTTCTTCCAGACTTTACAG 1140
DB 1268 CGTGGCTTAAGCATTAACAGCCAGTGTGATTCAGTTCGTTCTTCCAGACTTTACAG 1327
QY 1141 CAGATGCGTATTAACGCGCGCTGTCTTAATTAACGCGCGGAGAGTGAACCTGACC 1200
DB 1328 CAGATGCGTATTAACGCGCGCTGTCTTAATTAACGCGCGGAGAGTGAACCTGACC 1387
QY 1201 CTGGCAGATGTGAAGACGCAAAATGCGCAGGAGCATGTGTTGTCAGAGTTCAGCAA 1260
DB 1388 CTGGCAGATGTGAAGACGCAAAATGCGCAGGAGCATGTGTTGTCAGAGTTCAGCAA 1447
QY 1261 ATGCTGAATGAACGAGCATTAACCGCGCTGCAATTTG 1296
DB 1448 ATGCTGAATGAACGAGCATTAACCGCGCTGCAATTTG 1483

```

```

RESULT 14
US-08-910-798-1
; Sequence 1, Application US/08910798
; Patent No. 5876997
; GENERAL INFORMATION:
; APPLICANT: KRETZ
; TITLE OF INVENTION: NOVEL PHYTASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,798
; FILING DATE: August 13, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HAILE, PH.D., LISA A.
; REGISTRATION NUMBER: 38,347
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1272 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: PHYTASE
; FEATURE:
; NAME/KEY:
; LOCATION:
; US-08-910-798-1

```

Query Match 88.1%; Score 1153; DB 2; Length 1272;
 Best Local Similarity 94.1%; Pred. No. 0;
 Matches 1229; Conservative 0; Mismatches 26; Indels 51; Gaps 1;

```

QY 1 ATGAAAGCGATCTTAATCCCATTTTATCTCTTGATTCGCTTAACCCCGCAATCTGCA 60
DB 1 ATGAAAGCGATCTTAATCCCATTTTATCTCTTGATTCGCTTAACCCCGCAATCTGCA 60

```

```

QY 61 TTGCGCTCAGAGTGAAGCCGAGCTGAAAGCTGAAAGTGTGATTTGTCACTGCTATGCT 120
Db 61 TTGCGCTCAGAGTGAAGCCGAGCTGAAAGTGTGAAAGTGTGATTTGTCACTGCTATGCT 120
QY 121 GTGCGGTCTCAACCAAGGCCAGCAAGCTGATGACAGATGTCAAGCCAGAGCATATGGCA 180
Db 121 GTGCGGTCTCAACCAAGGCCAGCAAGCTGATGACAGATGTCAAGCCAGAGCATATGGCA 180
QY 181 ACCGTGGCCGGTAAACTGGGTGAAGCTGACACCGCGGGTGTGAGCTAAATGCGCTATCTC 240
Db 181 ACCGTGGCCGGTAAACTGGGTGAAGCTGACACCGCGGGTGTGAGCTAAATGCGCTATCTC 240
QY 241 GGAACATTAATGAGCTCAGAGCTGTGTGATGACCGAGATTTGCTGCTAAATGTGCTCCCG 300
Db 241 GGAACATTAATGAGCTCAGAGCTGTGTGATGACCGAGATTTGCTGCTAAATGTGCTCCCG 300
QY 301 CAGCTGTGCTCAGAGTGCAGATTAATGTGCTGATGTGACAGAGCTTAACCGGTAAACAGGCGAA 360
Db 301 CAGCTGTGCTCAGAGTGCAGATTAATGTGCTGATGTGACAGAGCTTAACCGGTAAACAGGCGAA 360
QY 361 GCGTTGCGCGCGGGGCTGACCTGACCTGACATTAACCGTACATACCCAGGCAAGATACG 420
Db 361 GCGTTGCGCGCGGGGCTGACCTGACCTGACATTAACCGTACATACCCAGGCAAGATACG 420
QY 421 TCCAGTCCCGGATCCGTTATTTATCTCTTAAACCTGAGCTTGGCCAACTGGATTAACGCG 480
Db 421 TCCAGTCCCGGATCCGTTATTTATCTCTTAAACCTGAGCTTGGCCAACTGGATTAACGCG 480
QY 481 AACGTGACTGACGCGATCTCTGAGAGAGGAGGAGGCTCAATTGCTGACTTTTACCGGGCAT 540
Db 481 AACGTGACTGACGCGATCTCTGAGAGAGGAGGAGGCTCAATTGCTGACTTTTACCGGGCAT 540
QY 541 TATCAACGCGGCTTTGCGCAACTGGAACGCGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600
Db 541 TATCAACGCGGCTTTGCGCAACTGGAACGCGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600
QY 601 CTTAAACGTGAGAAACAGAGAGAAAGCTTTCAATTACGAGGCAATTAACATCGGAATC 660
Db 601 CTTAAACGTGAGAAACAGAGAGAAAGCTTTCAATTACGAGGCAATTAACATCGGAATC 660
QY 661 AAGGTGAGGCGCGAGCTGTCTCATTTAACCGGTCGCTGATTAAGCTGCAATTAAGCTGAC 720
Db 661 AAGGTGAGGCGCGAGCTGTCTCATTTAACCGGTCGCTGATTAAGCTGCAATTAAGCTGAC 720
QY 721 GAGATATTTCTCTGCAACAGCAAGGAAATGCGGAGCCGAGGTGAGGAAAGATCAAC 780
Db 721 GAGATATTTCTCTGCAACAGCAAGGAAATGCGGAGCCGAGGTGAGGAAAGATCAAC 780
QY 781 GATTCAACACAGTGAACACCTTGCTAAGTTGCTAATACGCGCAATTGATTTGCTACAA 840
Db 781 GATTCAACACAGTGAACACCTTGCTAAGTTGCTAATACGCGCAATTGATTTGCTACAA 840
QY 841 CGCAGCGCAGAGGTTGCGCGAGCGCGCAACCGGTTATTAAGATTGATCAAGCAAGG 900
Db 841 CGCAGCGCAGAGGTTGCGCGAGCGCGCAACCGGTTATTAAGATTGATCAAGCAAGG 900
QY 901 TTGACGCGCCATCAACCGCAAAAACAGGCGTATGCTGACATTAACCACTTCAGTGTCT 960
Db 901 TTGACGCGCCATCAACCGCAAAAACAGGCGTATGCTGACATTAACCACTTCAGTGTCT 960
QY 961 TTTATTCGCGGACAGATTAATCTATCTGCGAAATCTCGCGCGCGCACTGGAGCTCACTGG 1020
Db 961 TTTATTCGCGGACAGATTAATCTATCTGCGAAATCTCGCGCGCGCACTGGAGCTCACTGG 1020
QY 1021 ACGCTTCCCGGTCAAGCGGATTAACAGCGCGCAGGAGGAGTGAATCTGATTTGAAGAGCTGG 1080
Db 1021 ACGCTTCCCGGTCAAGCGGATTAACAGCGCGCAGGAGGAGTGAATCTGATTTGAAGAGCTGG 1080
QY 1081 CGTGGGCTTAAGGATTAACAGCGAGTGAATTCAGGTTTGGCTGCTTCCAGACTTTACAG 1140
Db 1081 CGTGGGCTTAAGGATTAACAGCGAGTGAATTCAGGTTTGGCTGCTTCCAGACTTTACAG 1140
QY 1091 CGTGGGCTTAAGGATTAACAGCGAGTGAATTCAGGTTTGGCTGCTTCCAGACTTTACAG 1200
Db 1091 CGTGGGCTTAAGGATTAACAGCGAGTGAATTCAGGTTTGGCTGCTTCCAGACTTTACAG 1200

```

```

QY 1141 CAGATCGGTAAATAAACCCGCTGTCAATTAATACCGCCCGGAGAGTGAATCAACC 1200
Db 1090 CAGATCGGTAAATAAACCCGCTGTCAATTAATACCGCCCGGAGAGTGAATCAACC 1149
QY 1201 CTGCGAGATGTAAGAGAGCAAAATGCGAGGAGCATGTGTTGCTGGCAGGTTTACGCA 1260
Db 1150 CTGCGAGATGTAAGAGAGCAAAATGCGAGGAGCATGTGTTGCTGGCAGGTTTACGCA 1209
QY 1261 ATCGTAATGAAGACAGCATACCGGCGTGCAGTTTGAGATCTCATC 1306
Db 1210 ATCGTAATGAAGACAGCATACCGGCGTGCAGTTTGAGATCTCATC 1255

RESULT 15
US-09-489-039A-341
; Sequence 341, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NOCUTIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 341
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-341

Query Match 5.1%; Score 66.2; DB 3; Length 1266;
Best Local Similarity 49.2%; Pred. No. 9.4e-11;
Matches 204; Conservative 0; Mismatches 208; Indels 3; Gaps 1;

```

Search completed: June 14, 2006, 14:19:08
 Job time : 270.506 secs

THIS PAGE LEFT BLANK

Db 181 RQTAFRELEVLNFPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLLQQAQMPBEGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
Db 241 EIFLLQQAQMPBEGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
Qy 301 LTPHPQKQAYGVTLPFSVLFIAGHDTNLANLGALBELNMTLPGQPDNTPPGGELVBERW 360
Db 301 LTPHPQKQAYGVTLPFSVLFIAGHDTNLANLGALBELNMTLPGQPDNTPPGGELVBERW 360
Qy 361 RRLSDNSQMTQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Db 361 RRLSDNSQMTQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Qy 421 IYNBARIPACSLRSHHHHH 440
Db 421 IYNBARIPACSLRSHHHHH 440

RESULT 2

US-09-866-379-2
; Sequence 2, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER3170-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-09-866-379-2

Query Match 100.0%; Score 2302; DB 3; Length 440;

Best Local Similarity 100.0%; Pred. No. 2,2e-193;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKAILPFLSLILPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQLMQDVTTPDAMP 60
Db 1 MKAILPFLSLILPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQLMQDVTTPDAMP 60
Qy 61 TWPVKLGMLTPRGELIAYIGHYQORQLVADGLAKKGCPSQGVAIADVDERTRTKGE 120
Db 61 TWPVKLGMLTPRGELIAYIGHYQORQLVADGLAKKGCPSQGVAIADVDERTRTKGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTVGCQLDNANVTDAIISRAGSIAADFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTVGCQLDNANVTDAIISRAGSIAADFTGH 180
Qy 181 RQTAFRELEVLNFPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 RQTAFRELEVLNFPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy 241 EIFLLQQAQMPBEGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
Db 241 EIFLLQQAQMPBEGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
Qy 301 LTPHPQKQAYGVTLPFSVLFIAGHDTNLANLGALBELNMTLPGQPDNTPPGGELVBERW 360
Db 301 LTPHPQKQAYGVTLPFSVLFIAGHDTNLANLGALBELNMTLPGQPDNTPPGGELVBERW 360
Qy 361 RRLSDNSQMTQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Db 361 RRLSDNSQMTQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Qy 421 IYNBARIPACSLRSHHHHH 440
Db 421 IYNBARIPACSLRSHHHHH 440

RESULT 3

US-10-034-985-2
; Sequence 2, Application US/10034985
; Publication No. US20030049815A1
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/10/034,985
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US/09/580,515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-034-985-2

Query Match 100.0%; Score 2302; DB 4; Length 440;

Best Local Similarity 100.0%; Pred. No. 2,2e-193;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKAILPFLSLILPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQLMQDVTTPDAMP 60
Db 1 MKAILPFLSLILPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQLMQDVTTPDAMP 60
Qy 61 TWPVKLGMLTPRGELIAYIGHYQORQLVADGLAKKGCPSQGVAIADVDERTRTKGE 120
Db 61 TWPVKLGMLTPRGELIAYIGHYQORQLVADGLAKKGCPSQGVAIADVDERTRTKGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTVGCQLDNANVTDAIISRAGSIAADFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTVGCQLDNANVTDAIISRAGSIAADFTGH 180
Qy 181 RQTAFRELEVLNFPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 RQTAFRELEVLNFPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLLQQAQMPBEGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
Db 241 EIFLLQQAQMPBEGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
Qy 301 LTPHPQKQAYGVTLPFSVLFIAGHDTNLANLGALBELNMTLPGQPDNTPPGGELVBERW 360
Db 301 LTPHPQKQAYGVTLPFSVLFIAGHDTNLANLGALBELNMTLPGQPDNTPPGGELVBERW 360
Qy 361 RRLSDNSQMTQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Db 361 RRLSDNSQMTQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Db 361 RRLSDNSQWIGVSLVFQTLQQRMDKTPLSLNTPPGSEVKTLLAGCEBNAQWCSLAGFTQ 420
QY 421 IVNEARIPACSLRSHHHHH 440
Db 421 IVNEARIPACSLRSHHHHH 440

RESULT 4

US-10-430-356-2
; Sequence 2, Application US/10430356
; Publication No. US20030232041A1
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHRYASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/10/430,356
; PRIOR FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: US/09/580,515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-430-356-2

Query Match 100.0%; Score 2302; DB 4; Length 440;
Best Local Similarity 100.0%; Pred. No. 2.2e-193;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAILIPLSLILPLTPQSAFAOSEPELKESVIVSRHGVAPTKATQLMQDVTDPDMP 60
Db 1 MKAILIPLSLILPLTPQSAFAOSEPELKESVIVSRHGVAPTKATQLMQDVTDPDMP 60
QY 61 TWPVGLGWLTPRGSELIVLGHYQORLVADGLAKKCCPOSGQVAIIADVDERTRKTGE 120
Db 61 TWPVGLGWLTPRGSELIVLGHYQORLVADGLAKKCCPOSGQVAIIADVDERTRKTGE 120
QY 121 AFAAGLADPCAITVHTQADTSSPDPLFNPPLKTVGCQOLDNAVTTAIIISRAGGSIAADFTGH 180
Db 121 AFAAGLADPCAITVHTQADTSSPDPLFNPPLKTVGCQOLDNAVTTAIIISRAGGSIAADFTGH 180
QY 181 ROTAARELERLVANFPOSNLCIKREKODSCSLTQALPSELKVSADNVSLTGAIVSLASMLT 240
Db 181 ROTAARELERLVANFPOSNLCIKREKODSCSLTQALPSELKVSADNVSLTGAIVSLASMLT 240
QY 241 EIFLLQQAQMGPEPQWGRITDSHQNNTLLSHNAQFYLLQRTPEVARSRATPLDLIMAA 300
Db 241 EIFLLQQAQMGPEPQWGRITDSHQNNTLLSHNAQFYLLQRTPEVARSRATPLDLIMAA 300
QY 301 LTPHPPOKQAYGVTLPTSVLPIAGHDNTLANLGALBLNMTLLPQPDNTPPGSELVFERW 360
Db 301 LTPHPPOKQAYGVTLPTSVLPIAGHDNTLANLGALBLNMTLLPQPDNTPPGSELVFERW 360
QY 361 RRLSDNSQWIGVSLVFQTLQQRMDKTPLSLNTPPGSEVKTLLAGCEBNAQWCSLAGFTQ 420
Db 361 RRLSDNSQWIGVSLVFQTLQQRMDKTPLSLNTPPGSEVKTLLAGCEBNAQWCSLAGFTQ 420
QY 421 IVNEARIPACSLRSHHHHH 440
Db 421 IVNEARIPACSLRSHHHHH 440

RESULT 5

US-10-601-319-2
; Sequence 2, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O'Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT PHRYASES AND METHODS OF MAKING
; FILE REFERENCE: 09010-029011
; CURRENT APPLICATION NUMBER: US/10/601,319
; PRIOR FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-601-319-2

Query Match 100.0%; Score 2302; DB 4; Length 440;
Best Local Similarity 100.0%; Pred. No. 2.2e-193;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAILIPLSLILPLTPQSAFAOSEPELKESVIVSRHGVAPTKATQLMQDVTDPDMP 60
Db 1 MKAILIPLSLILPLTPQSAFAOSEPELKESVIVSRHGVAPTKATQLMQDVTDPDMP 60
QY 61 TWPVGLGWLTPRGSELIVLGHYQORLVADGLAKKCCPOSGQVAIIADVDERTRKTGE 120
Db 61 TWPVGLGWLTPRGSELIVLGHYQORLVADGLAKKCCPOSGQVAIIADVDERTRKTGE 120
QY 121 AFAAGLADPCAITVHTQADTSSPDPLFNPPLKTVGCQOLDNAVTTAIIISRAGGSIAADFTGH 180
Db 121 AFAAGLADPCAITVHTQADTSSPDPLFNPPLKTVGCQOLDNAVTTAIIISRAGGSIAADFTGH 180
QY 181 ROTAARELERLVANFPOSNLCIKREKODSCSLTQALPSELKVSADNVSLTGAIVSLASMLT 240
Db 181 ROTAARELERLVANFPOSNLCIKREKODSCSLTQALPSELKVSADNVSLTGAIVSLASMLT 240
QY 241 EIFLLQQAQMGPEPQWGRITDSHQNNTLLSHNAQFYLLQRTPEVARSRATPLDLIMAA 300
Db 241 EIFLLQQAQMGPEPQWGRITDSHQNNTLLSHNAQFYLLQRTPEVARSRATPLDLIMAA 300
QY 301 LTPHPPOKQAYGVTLPTSVLPIAGHDNTLANLGALBLNMTLLPQPDNTPPGSELVFERW 360
Db 301 LTPHPPOKQAYGVTLPTSVLPIAGHDNTLANLGALBLNMTLLPQPDNTPPGSELVFERW 360
QY 361 RRLSDNSQWIGVSLVFQTLQQRMDKTPLSLNTPPGSEVKTLLAGCEBNAQWCSLAGFTQ 420
Db 361 RRLSDNSQWIGVSLVFQTLQQRMDKTPLSLNTPPGSEVKTLLAGCEBNAQWCSLAGFTQ 420
QY 421 IVNEARIPACSLRSHHHHH 440
Db 421 IVNEARIPACSLRSHHHHH 440

```
RESULT 6
US-10-933-115-2
; Sequence 2, Application US/10933115
; Publication No. US20050281792A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O'Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES
; FILE REFERENCE: 564462001822
; CURRENT APPLICATION NUMBER: US/10/933,115
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRN
; ORGANISM: Escherichia coli
US-10-933-115-2

Query Match      100.0%; Score 2302; DB 5; Length 440;
Best Local Similarity 100.0%; Pred. No. 2,2e-193;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAILIPFLSLILPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
DB 1 MKAILIPFLSLILPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
QY 61 TWPVKLGMLTPRGSELAYIGHYQORQRLVADGLAKKGCPOSGQVAIIADVDETRKTGE 120
DB 61 TWPVKLGMLTPRGSELAYIGHYQORQRLVADGLAKKGCPOSGQVAIIADVDETRKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVQCLDNANVTDAIISRAGSIAADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVQCLDNANVTDAIISRAGSIAADFTGH 180
QY 181 RQTAFRELEVLNFPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 RQTAFRELEVLNFPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLOAQOQMPBEGWRITDSHOWNTLSLHNAQFYLLQRTPEVASRATPLDLIMAA 300
DB 241 EIFLLOAQOQMPBEGWRITDSHOWNTLSLHNAQFYLLQRTPEVASRATPLDLIMAA 300
QY 301 LTPHPQKQAYGVTLPTSVLFIAGHDNTLANLGALBELNMTLPGQPDNTPPGSELVFERW 360
DB 301 LTPHPQKQAYGVTLPTSVLFIAGHDNTLANLGALBELNMTLPGQPDNTPPGSELVFERW 360
QY 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGSEVYKTLACGEBERNAQGMCSLAGFTQ 420
DB 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGSEVYKTLACGEBERNAQGMCSLAGFTQ 420
QY 421 IVNEARIPACSLRSHHHHH 440
DB 421 IVNEARIPACSLRSHHHHH 440
```

```
RESULT 7
US-09-866-379-8
; Sequence 8, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 432
; TYPE: PRN
; ORGANISM: Escherichia coli
US-09-866-379-8
```

```
Query Match      97.1%; Score 2235; DB 3; Length 432;
Best Local Similarity 99.5%; Pred. No. 1,7e-187;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKAILIPFLSLILPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
DB 1 MKAILIPFLSLILPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
QY 61 TWPVKLGMLTPRGSELAYIGHYQORQRLVADGLAKKGCPOSGQVAIIADVDETRKTGE 120
DB 61 TWPVKLGMLTPRGSELAYIGHYQORQRLVADGLAKKGCPOSGQVAIIADVDETRKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVQCLDNANVTDAIISRAGSIAADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVQCLDNANVTDAIISRAGSIAADFTGH 180
QY 181 RQTAFRELEVLNFPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 RQTAFRELEVLNFPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLOAQOQMPBEGWRITDSHOWNTLSLHNAQFYLLQRTPEVASRATPLDLIMAA 300
DB 241 EIFLLOAQOQMPBEGWRITDSHOWNTLSLHNAQFYLLQRTPEVASRATPLDLIMAA 300
QY 301 LTPHPQKQAYGVTLPTSVLFIAGHDNTLANLGALBELNMTLPGQPDNTPPGSELVFERW 360
DB 301 LTPHPQKQAYGVTLPTSVLFIAGHDNTLANLGALBELNMTLPGQPDNTPPGSELVFERW 360
QY 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGSEVYKTLACGEBERNAQGMCSLAGFTQ 420
DB 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGSEVYKTLACGEBERNAQGMCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432
```

RESULT 8

```

US-10-156-660-4
; Sequence 4, Application US/10156660
; Publication No. US20030103958A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson R.
; APPLICANT: Garrett, James B.
; APPLICANT: O'Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: PHYLASES, NUCLEIC ACIDS ENCODING THEM
; FILE REFERENCE: 09010-029007
; CURRENT APPLICATION NUMBER: US/10/156,660
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 432
; TYPE: prt
; ORGANISM: Escherichia coli
US-10-156-660-4

```

```

Query Match      97.1% Score 2235; DB 4; Length 432;
Best Local Similarity 99.5%; Pred. No. 1.7e-187;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 MKAIIIPFLSLIPLTPGSAFAQSEBELKLSVIVSRHGVAAPTKATQMLMDVTPDAMP 60
DB 1 MKAIIIPFLSLIPLTPGSAFAQSEBELKLSVIVSRHGVAAPTKATQMLMDVTPDAMP 60
QY 61 TWPVGLGMLTRGGELIAYLGHYORQRLVADGLAKKGCPOGGOVAIADVDERTKTGE 120
DB 61 TWPVGLGMLTRGGELIAYLGHYORQRLVADGLAKKGCPOGGOVAIADVDERTKTGE 120
QY 121 APAAGLAPDCAITVHTQADTSSPPLFNPPLKTGVCOLDNANVTDAIISRAGSIAIDFTGH 180
DB 121 APAAGLAPDCAITVHTQADTSSPPLFNPPLKTGVCOLDNANVTDAIISRAGSIAIDFTGH 180
QY 181 RQTAARELERVLPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 RQTAARELERVLPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLQQAQGMPEPBGKRTTDSHQNNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIKTA 300
DB 241 EIFLLQQAQGMPEPBGKRTTDSHQNNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIKTA 300
QY 301 LTPHPPOKQAVGVTLPSTVLPFAGHDITNLNANGALLETNMTLPGQPDNTPPGSELVPERW 360
DB 301 LTPHPPOKQAVGVTLPSTVLPFAGHDITNLNANGALLETNMTLPGQPDNTPPGSELVPERW 360
QY 361 RRLSDNSQMIQVSLVFQTLQQRMDTLPISLNTPESEVKLTLLAGCERNAQGCISLAGFTQ 420
DB 361 RRLSDNSQMIQVSLVFQTLQQRMDTLPISLNTPESEVKLTLLAGCERNAQGCISLAGFTQ 420
QY 421 IYNEKRIAPCSL 432
DB 421 IYNEKRIAPCSL 432

```

RESULT 9

```

US-10-282-122A-43351
; Sequence 43351, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patencin version 3.1
; SEQ ID NO 43351
; LENGTH: 432
; TYPE: prt
; ORGANISM: Escherichia coli
US-10-282-122A-43351

```

```

Query Match      97.1% Score 2235; DB 4; Length 432;
Best Local Similarity 99.5%; Pred. No. 1.7e-187;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 MKAIIIPFLSLIPLTPGSAFAQSEBELKLSVIVSRHGVAAPTKATQMLMDVTPDAMP 60
DB 1 MKAIIIPFLSLIPLTPGSAFAQSEBELKLSVIVSRHGVAAPTKATQMLMDVTPDAMP 60
QY 61 TWPVGLGMLTRGGELIAYLGHYORQRLVADGLAKKGCPOGGOVAIADVDERTKTGE 120
DB 61 TWPVGLGMLTRGGELIAYLGHYORQRLVADGLAKKGCPOGGOVAIADVDERTKTGE 120
QY 121 APAAGLAPDCAITVHTQADTSSPPLFNPPLKTGVCOLDNANVTDAIISRAGSIAIDFTGH 180
DB 121 APAAGLAPDCAITVHTQADTSSPPLFNPPLKTGVCOLDNANVTDAIISRAGSIAIDFTGH 180
QY 181 RQTAARELERVLPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 RQTAARELERVLPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLQQAQGMPEPBGKRTTDSHQNNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIKTA 300
DB 241 EIFLLQQAQGMPEPBGKRTTDSHQNNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIKTA 300

```

QY 301 LTPHPKOKAYGVTLPSTVLFIAHDITNLANIAGALBLNMTLPQOPDNTPPGSELVFERW 360
DB 301 LTPHPKOKAYGVTLPSTVLFIAHDITNLANIAGALBLNMTLPQOPDNTPPGSELVFERW 360
QY 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCISLAGFTQ 420
DB 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCISLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 10

US-10-601-319-8
; Sequence 8, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O'Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT PHRYASASES AND METHODS OF MAKING
; TITLE OF INVENTION: AND USING THEM
; FILE REFERENCE: 09010-029011
; CURRENT FILING DATE: US/10/601,319
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-601-319-8

Query Match 97.1%; Score 2235; DB 4; Length 432;

Best Local Similarity 99.5%; Pred. No. 1.7e-187;

Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKAILIPFLSLILPTPOSAPFAOSEPELKESVIVSRHGVARPTKATQLMQDVTTPDAMP 60
DB 1 MKAILIPFLSLILPTPOSAPFAOSEPELKESVIVSRHGVARPTKATQLMQDVTTPDAMP 60
QY 61 TWPVKGWLTTPRGSELIAVIGHYQORLVADGLAKKGCPOSGOVAIADVDERTRKTGE 120
DB 61 TWPVKGWLTTPRGSELIAVIGHYQORLVADGLAKKGCPOSGOVAIADVDERTRKTGE 120
QY 121 APAAGLAPDCAITVHTQADTSSPDPLFNPDKTGVCOQDNNANVTDAIISRAGGSIAADFTG 180
DB 121 APAAGLAPDCAITVHTQADTSSPDPLFNPDKTGVCOQDNNANVTDAIISRAGGSIAADFTG 180
QY 181 ROTAARELERVLPNPOSMLCKREKODESCSLTQALPSELKVSADNVSITGAVSLASMLT 240
DB 181 ROTAARELERVLPNPOSMLCKREKODESCSLTQALPSELKVSADNVSITGAVSLASMLT 240
QY 241 EIFLLQQAQMPBEQWGRITDSHQNNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
DB 241 EIFLLQQAQMPBEQWGRITDSHQNNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIMAA 300

DB 241 EIFLLQQAQMPBEQWGRITDSHQNNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIKTA 300
QY 301 LTPHPKOKAYGVTLPSTVLFIAHDITNLANIAGALBLNMTLPQOPDNTPPGSELVFERW 360
DB 301 LTPHPKOKAYGVTLPSTVLFIAHDITNLANIAGALBLNMTLPQOPDNTPPGSELVFERW 360
QY 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCISLAGFTQ 420
DB 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCISLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 11

US-10-472-317-41
; Sequence 41, Application US/10472317
; Publication No. US20040185562A1
; GENERAL INFORMATION:
; APPLICANT: Cargill, Incorporated
; TITLE OF INVENTION: Myo-Inositol Oxygenases
; FILE REFERENCE: 10829/003US1
; CURRENT FILING DATE: US/10/472,317
; PRIOR APPLICATION NUMBER: PCT/US02/08404
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: US 60/277,148
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-472-317-41

Query Match 97.1%; Score 2235; DB 4; Length 432;

Best Local Similarity 99.5%; Pred. No. 1.7e-187;

Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKAILIPFLSLILPTPOSAPFAOSEPELKESVIVSRHGVARPTKATQLMQDVTTPDAMP 60
DB 1 MKAILIPFLSLILPTPOSAPFAOSEPELKESVIVSRHGVARPTKATQLMQDVTTPDAMP 60
QY 61 TWPVKGWLTTPRGSELIAVIGHYQORLVADGLAKKGCPOSGOVAIADVDERTRKTGE 120
DB 61 TWPVKGWLTTPRGSELIAVIGHYQORLVADGLAKKGCPOSGOVAIADVDERTRKTGE 120
QY 121 APAAGLAPDCAITVHTQADTSSPDPLFNPDKTGVCOQDNNANVTDAIISRAGGSIAADFTG 180
DB 121 APAAGLAPDCAITVHTQADTSSPDPLFNPDKTGVCOQDNNANVTDAIISRAGGSIAADFTG 180
QY 181 ROTAARELERVLPNPOSMLCKREKODESCSLTQALPSELKVSADNVSITGAVSLASMLT 240
DB 181 ROTAARELERVLPNPOSMLCKREKODESCSLTQALPSELKVSADNVSITGAVSLASMLT 240
QY 241 EIFLLQQAQMPBEQWGRITDSHQNNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
DB 241 EIFLLQQAQMPBEQWGRITDSHQNNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIKTA 300
QY 301 LTPHPKOKAYGVTLPSTVLFIAHDITNLANIAGALBLNMTLPQOPDNTPPGSELVFERW 360
DB 301 LTPHPKOKAYGVTLPSTVLFIAHDITNLANIAGALBLNMTLPQOPDNTPPGSELVFERW 360
QY 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCISLAGFTQ 420
DB 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCISLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

```
RESULT 12
US-10-933-115-8
; Sequence 8, Application US/1093115
; Publication No. US20050281792A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYLASES AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 564462001822
; CURRENT APPLICATION NUMBER: US/10/933,115
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-933-115-8

Query Match      97.1%; Score 2235; DB 5; Length 432;
Beet Local Similarity 99.5%; Pred. No. 1.7e-187;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKAILIPFLSLILPTLPSAFAQSEBELKESVIVSRHGVRAPTKATQMDVTPDMP 60
DB 1 MKAILIPFLSLILPTLPSAFAQSEBELKESVIVSRHGVRAPTKATQMDVTPDMP 60
QY 61 TWPVGLGWLTPRGSELIAVLGHYQORQLVADGLAKKGCPOSGQVAIIADVDETRKTGE 120
DB 61 TWPVGLGWLTPRGSELIAVLGHYQORQLVADGLAKKGCPOSGQVAIIADVDETRKTGE 120
QY 121 APAAGLADPCAITVHTQADTSSPDLFNPPLKTGVCQOLDNAVNTDAIISRAGGSIAADFTGH 180
DB 121 APAAGLADPCAITVHTQADTSSPDLFNPPLKTGVCQOLDNAVNTDAIISRAGGSIAADFTGH 180
QY 181 ROTAFAELERLVANPQSNLCLKREKODSCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFAELERLVANPQSNLCLKREKODSCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLOAQOQMPBEQWGRITDSSHQWNTLLSIHNAQFYLLQRTPEVARSRAATPLDLIYMAA 300
DB 241 EIFLLOAQOQMPBEQWGRITDSSHQWNTLLSIHNAQFYLLQRTPEVARSRAATPLDLIYMAA 300
QY 301 LTPHPQOQAGVTLPTSVLPIAGHDNLNLAGALEINMTLPGQPDNTTPGGEIVFERW 360
DB 301 LTPHPQOQAGVTLPTSVLPIAGHDNLNLAGALEINMTLPGQPDNTTPGGEIVFERW 360
QY 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPGGEVLTLAGCERNAAQCMCSLAGFTQ 420
DB 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPGGEVLTLAGCERNAAQCMCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432
```

```
RESULT 13
US-11-056-354-4
; Sequence 4, Application US/11056354
; Publication No. US20050246780A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson R.
; APPLICANT: Garrett, James B.
; APPLICANT: O'DONOGHUE, Eileen
; APPLICANT: MATHUR, ERIC J.
; TITLE OF INVENTION: PHYLASES, NUCLEIC ACIDS ENCODING THEM AND METHODS
; TITLE OF INVENTION: FOR MAKING AND USING THEM
; FILE REFERENCE: 564462001803/D1370-9C1
; CURRENT APPLICATION NUMBER: US/11/056,354
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 10/156,660
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli appa phytase
US-11-056-354-4

Query Match      97.1%; Score 2235; DB 6; Length 432;
Beet Local Similarity 99.5%; Pred. No. 1.7e-187;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKAILIPFLSLILPTLPSAFAQSEBELKESVIVSRHGVRAPTKATQMDVTPDMP 60
DB 1 MKAILIPFLSLILPTLPSAFAQSEBELKESVIVSRHGVRAPTKATQMDVTPDMP 60
QY 61 TWPVGLGWLTPRGSELIAVLGHYQORQLVADGLAKKGCPOSGQVAIIADVDETRKTGE 120
DB 61 TWPVGLGWLTPRGSELIAVLGHYQORQLVADGLAKKGCPOSGQVAIIADVDETRKTGE 120
QY 121 APAAGLADPCAITVHTQADTSSPDLFNPPLKTGVCQOLDNAVNTDAIISRAGGSIAADFTGH 180
DB 121 APAAGLADPCAITVHTQADTSSPDLFNPPLKTGVCQOLDNAVNTDAIISRAGGSIAADFTGH 180
QY 181 ROTAFAELERLVANPQSNLCLKREKODSCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFAELERLVANPQSNLCLKREKODSCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLOAQOQMPBEQWGRITDSSHQWNTLLSIHNAQFYLLQRTPEVARSRAATPLDLIYMAA 300
DB 241 EIFLLOAQOQMPBEQWGRITDSSHQWNTLLSIHNAQFYLLQRTPEVARSRAATPLDLIYMAA 300
QY 301 LTPHPQOQAGVTLPTSVLPIAGHDNLNLAGALEINMTLPGQPDNTTPGGEIVFERW 360
DB 301 LTPHPQOQAGVTLPTSVLPIAGHDNLNLAGALEINMTLPGQPDNTTPGGEIVFERW 360
QY 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPGGEVLTLAGCERNAAQCMCSLAGFTQ 420
DB 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPGGEVLTLAGCERNAAQCMCSLAGFTQ 420
```

QY 421 IYNEARIPACSL 432
Db 421 IYNEARIPACSL 432

RESULT 14

US-11-018-709-1
; Sequence 1, Application US/11018709
; Publication No. US20050095691A1
; GENERAL INFORMATION:
; APPLICANT: Lei, Xiangen
; TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYTASE
; FILE REFERENCE: 19603/4031
; CURRENT APPLICATION NUMBER: US/11/018,709
; PRIOR FILING DATE: 2004-12-20 715,477
; PRIOR APPLICATION NUMBER: US/09/715,477
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,179
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-11-018-709-1

Query Match 96.9%; Score 2230; DB 6; Length 432;
Best Local Similarity 99.3%; Pred. No. 4,6e-187;
Matches 429; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKAILIPFLSLILPLTPQSAFAOSEPELKESVIVSRHGVRAPTKATQLMQDVTDPAMP 60
Db 1 MKAILIPFLSLILPLTPQSAFAOSEPELKESVIVSRHGVRAPTKATQLMQDVTDPAMP 60
QY 61 TWPKLGMFLTPRGSELIAVIGHYQORQLVADGLAKKGCPOGVAIIADVDETRKTGE 120
Db 61 TWPKLGMFLTPRGSELIAVIGHYQORQLVADGLAKKGCPOGVAIIADVDETRKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNNAVTDAILSRAGGSIAADFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNNAVTDAILSRAGGSIAADFTGH 180
QY 181 RQTAFRELEVLNFPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 RQTAFRELEVLNFPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLQQAQGMPEBGMGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPPLDLIMAA 300
Db 241 EIFLLQQAQGMPEBGMGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPPLDLIMAA 300
QY 301 LTPHPQKQAYGVTLPSTVLFIAGHDTNLNLGALBELNMTLPGQPDNTPPGGELVPERW 360
Db 301 LTPHPQKQAYGVTLPSTVLFIAGHDTNLNLGALBELNMTLPGQPDNTPPGGELVPERW 360
QY 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEBNAQGMCSLAGFTQ 420
Db 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEBNAQGMCSLAGFTQ 420
QY 421 IYNEARIPACSL 432
Db 421 IYNEARIPACSL 432

RESULT 15

US-10-284-962-3
; Sequence 3, Application US/10284962
; Publication No. US20030206913A1
; GENERAL INFORMATION:
; APPLICANT: Webel, Douglas M.
; APPLICANT: Orr, Donald E.
; APPLICANT: Ruch, Frank E.
; TITLE OF INVENTION: PHYTASE-CONTAINING ANIMAL FOOD AND METHOD

FILE REFERENCE: 834460-71725
; CURRENT APPLICATION NUMBER: US/10/284,962
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/335,303
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-284-962-3

Query Match 96.5%; Score 2221; DB 4; Length 432;
Best Local Similarity 99.1%; Pred. No. 2,8e-186;
Matches 428; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MKAILIPFLSLILPLTPQSAFAOSEPELKESVIVSRHGVRAPTKATQLMQDVTDPAMP 60
Db 1 MKAILIPFLSLILPLTPQSAFAOSEPELKESVIVSRHGVRAPTKATQLMQDVTDPAMP 60
QY 61 TWPKLGMFLTPRGSELIAVIGHYQORQLVADGLAKKGCPOGVAIIADVDETRKTGE 120
Db 61 TWPKLGMFLTPRGSELIAVIGHYQORQLVADGLAKKGCPOGVAIIADVDETRKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNNAVTDAILSRAGGSIAADFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNNAVTDAILSRAGGSIAADFTGH 180
QY 181 RQTAFRELEVLNFPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 RQTAFRELEVLNFPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLQQAQGMPEBGMGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPPLDLIMAA 300
Db 241 EIFLLQQAQGMPEBGMGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPPLDLIMAA 300
QY 301 LTPHPQKQAYGVTLPSTVLFIAGHDTNLNLGALBELNMTLPGQPDNTPPGGELVPERW 360
Db 301 LTPHPQKQAYGVTLPSTVLFIAGHDTNLNLGALBELNMTLPGQPDNTPPGGELVPERW 360
QY 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEBNAQGMCSLAGFTQ 420
Db 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEBNAQGMCSLAGFTQ 420
QY 421 IYNEARIPACSL 432
Db 421 IYNEARIPACSL 432

Search completed: June 13, 2006, 10:50:37
Job time : 145.303 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2006, 10:28:49 ; Search time 36.3303 Seconds
(without alignments)
1060.093 Million cell updates/sec

Title: US-10-601-319-2
Perfect score: 2302
Sequence: 1 MKALIFPLSLILPLTPQSA.....ITNEARIPACSLRSHHHHH 440

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*
1: /EMC_Celestra_SIDS3/ptodata/2/1aa/5_COMB pep:*
2: /EMC_Celestra_SIDS3/ptodata/2/1aa/6_COMB pep:*
3: /EMC_Celestra_SIDS3/ptodata/2/1aa/7_COMB pep:*
4: /EMC_Celestra_SIDS3/ptodata/2/1aa/H_COMB pep:*
5: /EMC_Celestra_SIDS3/ptodata/2/1aa/PCYUS_COMB pep:*
6: /EMC_Celestra_SIDS3/ptodata/2/1aa/RE_COMB pep:*
7: /EMC_Celestra_SIDS3/ptodata/2/1aa/backfill1est.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2302	100.0	440	2	US-09-259-214-2
2	2302	100.0	440	2	US-09-318-528-2
3	2302	100.0	440	2	US-09-291-931-2
4	2302	100.0	440	2	US-09-580-515-2
5	2302	100.0	440	2	US-09-866-379-2
6	2235	97.1	432	2	US-09-866-379-8
7	2230	96.9	432	2	US-09-715-477-1
8	2231	96.5	432	2	US-10-266-041A-1
9	2221	96.5	433	2	US-09-540-149A-1
10	2210	96.0	432	2	US-09-715-477-3
11	2194.5	95.3	423	1	US-08-910-798-2
12	2141	93.0	430	2	US-09-866-379-10
13	528	22.9	422	2	US-09-489-039A-7512
14	376	16.3	521	2	US-09-489-039A-13501
15	119	5.2	439	2	US-09-044-718-3
16	119	5.2	439	2	US-10-062-848-3
17	119	5.2	440	2	US-09-684-855-106
18	119	5.2	440	2	US-09-684-855-128
19	119	5.2	440	2	US-09-684-855-151
20	119	5.2	440	2	US-09-488-265B-6
21	119	5.2	449	2	US-09-044-718-12
22	119	5.2	449	2	US-10-062-848-12
23	119	5.2	465	2	US-08-868-435-33
24	119	5.2	465	2	US-08-744-231-33
25	119	5.2	465	2	US-09-044-718-78
26	119	5.2	465	2	US-09-636-499-6

27	119	5.2	465	2	US-09-273-871A-8	Sequence 8, Appl1
28	119	5.2	465	2	US-10-083-452-8	Sequence 8, Appl1
29	119	5.2	465	2	US-09-635-504-33	Sequence 33, Appl1
30	119	5.2	465	2	US-10-062-848-78	Sequence 78, Appl1
31	119	5.2	465	3	US-10-229-358-6	Sequence 6, Appl1
32	118	5.1	440	2	US-09-684-855-107	Sequence 107, App
33	118	5.1	440	2	US-09-684-855-129	Sequence 129, App
34	118	5.1	440	2	US-09-684-855-152	Sequence 152, App
35	118	5.1	440	2	US-09-488-265B-7	Sequence 7, Appl1
36	117	5.1	386	1	US-08-758-213-1	Sequence 48, Appl1
37	117	5.1	386	1	US-08-692-787-48	Sequence 48, Appl1
38	117	5.1	386	2	US-09-097-199-48	Sequence 6022, App
39	117	5.1	386	2	US-09-949-016-6022	Sequence 2, Appl1
40	117	5.1	515	1	US-09-146-283-2	Sequence 2, Appl1
41	117	5.1	515	2	US-08-579-823A-2	Sequence 7, Appl1
42	117	5.1	515	2	US-09-344-195-2	Sequence 2, Appl1
43	114.5	5.0	326	2	US-09-270-767-41623	Sequence 41623, A
44	114.5	5.0	585	3	US-10-144-678A-1020	Sequence 1020, App
45	111	4.8	440	2	US-09-684-855-108	Sequence 108, App

ALIGNMENTS

RESULT 1						
US-09-259-214-2						
Sequence 2, Application US/09259214A						
Patent No. 6110719						
GENERAL INFORMATION:						
APPLICANT: Kreitz, Keith						
TITLE OF INVENTION: NOVEL PHYTASE						
FILE REFERENCE: DIVER370-1						
CURRENT APPLICATION NUMBER: US/09/259,214A						
CURRENT FILING DATE: 1999-03-01						
EARLIER APPLICATION NUMBER: 08/910,798						
EARLIER FILING DATE: 1997-08-13						
NUMBER OF SEQ ID NOS: 4						
SOFTWARE: FASTSEQ for Windows Version 4.0						
SEQ ID NO 2						
LENGTH: 440						
TYPE: PRT						
ORGANISM: Escherichia coli						
US-09-259-214-2						
Query Match						
Best Local Similarity 100.0%; Score 2302; DB 2; Length 440;						
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
QY	1					
DB	1					
QY	61					
DB	61					
QY	121					
DB	121					
QY	181					
DB	181					
QY	241					
DB	241					
QY	301					
DB	301					
QY	361					
DB	361					

```
Db      361 RRLSDNSQMWIOVSLVFQTLQMRDXTPLSLNTPPGEVLTLAGCEERNAQMGCSLAGFTQ 420
Qy      421 IVNEARIPACSLRSHHHHH 440
Db      421 IVNEARIPACSLRSHHHHH 440

RESULT 2
US-09-318-528-2
; Sequence 2, Application US/09318528
; Patent No. 6183740
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/318,528
; CURRENT FILING DATE: 1999-05-25
; EARLIER APPLICATION NUMBER: 09/291,931
; EARLIER FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-318-528-2

Query Match      100.0%; Score 2302; DB 2; Length 440;
Best Local Similarity 100.0%; Pred. No. 3.9e-234;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MKAILIPFLSLILPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTTPDAMP 60
Db      1 MKAILIPFLSLILPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTTPDAMP 60
Qy      61 TWPKLGMVLPGRGELIAYIGHYQORLVADGLAKKGCPOSGQVAIIADVDERTRKTGE 120
Db      61 TWPKLGMVLPGRGELIAYIGHYQORLVADGLAKKGCPOSGQVAIIADVDERTRKTGE 120
Qy      121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOQDNNANVTDAIISRAGSIADFTGH 180
Db      121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOQDNNANVTDAIISRAGSIADFTGH 180
Qy      121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOQDNNANVTDAIISRAGSIADFTGH 180
Db      121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOQDNNANVTDAIISRAGSIADFTGH 180
Qy      181 RQTAFRELERVLNFPQSNLCIKREKQDSCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db      181 RQTAFRELERVLNFPQSNLCIKREKQDSCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy      181 RQTAFRELERVLNFPQSNLCIKREKQDSCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db      181 RQTAFRELERVLNFPQSNLCIKREKQDSCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy      241 EIFLLOQAQGMPEBGMGRITDISHQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
Db      241 EIFLLOQAQGMPEBGMGRITDISHQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
Qy      301 LTPHPPOQAAYGVTLPTSVLPIAGHDNTLANLGALBELANMTLPQOPDNTPPGSELVEFERW 360
Db      301 LTPHPPOQAAYGVTLPTSVLPIAGHDNTLANLGALBELANMTLPQOPDNTPPGSELVEFERW 360
Qy      361 RRLSDNSQMWIOVSLVFQTLQMRDXTPLSLNTPPGEVLTLAGCEERNAQMGCSLAGFTQ 420
Db      361 RRLSDNSQMWIOVSLVFQTLQMRDXTPLSLNTPPGEVLTLAGCEERNAQMGCSLAGFTQ 420
Qy      421 IVNEARIPACSLRSHHHHH 440
Db      421 IVNEARIPACSLRSHHHHH 440

RESULT 3
US-09-291-931-2
; Sequence 2, Application US/09291931A
; Patent No. 6190897
```

```
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/291,931A
; CURRENT FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-291-931-2

Query Match      100.0%; Score 2302; DB 2; Length 440;
Best Local Similarity 100.0%; Pred. No. 3.9e-234;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MKAILIPFLSLILPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTTPDAMP 60
Db      1 MKAILIPFLSLILPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTTPDAMP 60
Qy      61 TWPKLGMVLPGRGELIAYIGHYQORLVADGLAKKGCPOSGQVAIIADVDERTRKTGE 120
Db      61 TWPKLGMVLPGRGELIAYIGHYQORLVADGLAKKGCPOSGQVAIIADVDERTRKTGE 120
Qy      121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOQDNNANVTDAIISRAGSIADFTGH 180
Db      121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOQDNNANVTDAIISRAGSIADFTGH 180
Qy      121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOQDNNANVTDAIISRAGSIADFTGH 180
Db      121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOQDNNANVTDAIISRAGSIADFTGH 180
Qy      181 RQTAFRELERVLNFPQSNLCIKREKQDSCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db      181 RQTAFRELERVLNFPQSNLCIKREKQDSCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy      181 RQTAFRELERVLNFPQSNLCIKREKQDSCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db      181 RQTAFRELERVLNFPQSNLCIKREKQDSCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy      241 EIFLLOQAQGMPEBGMGRITDISHQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
Db      241 EIFLLOQAQGMPEBGMGRITDISHQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
Qy      301 LTPHPPOQAAYGVTLPTSVLPIAGHDNTLANLGALBELANMTLPQOPDNTPPGSELVEFERW 360
Db      301 LTPHPPOQAAYGVTLPTSVLPIAGHDNTLANLGALBELANMTLPQOPDNTPPGSELVEFERW 360
Qy      361 RRLSDNSQMWIOVSLVFQTLQMRDXTPLSLNTPPGEVLTLAGCEERNAQMGCSLAGFTQ 420
Db      361 RRLSDNSQMWIOVSLVFQTLQMRDXTPLSLNTPPGEVLTLAGCEERNAQMGCSLAGFTQ 420
Qy      421 IVNEARIPACSLRSHHHHH 440
Db      421 IVNEARIPACSLRSHHHHH 440

RESULT 4
US-09-580-515-2
; Sequence 2, Application US/09580515
; Patent No. 6720014
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/580,515
; CURRENT FILING DATE: 1999-05-25
; EARLIER APPLICATION NUMBER: 09/318,528
; EARLIER FILING DATE: 1999-05-25
; EARLIER APPLICATION NUMBER: 09/291,931
; EARLIER FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
```

NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 440
TYPE: PRT
ORGANISM: Escherichia coli
US-09-580-515-2

Query Match 100.0%; Score 2302; DB 2; Length 440;
Best Local Similarity 100.0%; Pred. No. 3.9e-234;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MKAIIIPFLSLIPLTPPSAFASPELKLSEVIVSRHGVRAPRTAKTQMDVTPDAMP 60
DB 1 MKAIIIPFLSLIPLTPPSAFASPELKLSEVIVSRHGVRAPRTAKTQMDVTPDAMP 60
QY 61 TWPVGLGWLTPRGELIAYLGHYQORLVADGLAKKCCPOSGQVAIIADVDETRKTGE 120
DB 61 TWPVGLGWLTPRGELIAYLGHYQORLVADGLAKKCCPOSGQVAIIADVDETRKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPDKTGVCQLDNANVTDAIISRAGSIADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPDKTGVCQLDNANVTDAIISRAGSIADFTGH 180
QY 181 ROTARELERVLPQSNMCLKREKODESCSLTQALPSBLKVSADNVSLTGAVSLASMLT 240
DB 181 ROTARELERVLPQSNMCLKREKODESCSLTQALPSBLKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLQQAQGMPEPMGRITDISHQWNTLLSLHNAQFYLLQRTPEVARSRATPLDLINMA 300
DB 241 EIFLLQQAQGMPEPMGRITDISHQWNTLLSLHNAQFYLLQRTPEVARSRATPLDLINMA 300
QY 301 LTPHPPOQAQVGTLPSTVLPFIAGHDITMLANIGALBELNMTLPQGPDMTPPGSELVFERW 360
DB 301 LTPHPPOQAQVGTLPSTVLPFIAGHDITMLANIGALBELNMTLPQGPDMTPPGSELVFERW 360
QY 361 RLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPPEVGLTLAGCERNAGCISLAGFTQ 420
DB 361 RLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPPEVGLTLAGCERNAGCISLAGFTQ 420
QY 421 IVNEARIPACSLRSHHHHH 440
DB 421 IVNEARIPACSLRSHHHHH 440
```

RESULT 5
US-09-866-379-2
Sequence 2, Application US/09866379
Patent No. 6855365

GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: SHORT, Jay
APPLICANT: KRETZ, Keith
APPLICANT: GRAY, Kevin
APPLICANT: BARTON, Nelson
APPLICANT: GARRETT, James
APPLICANT: O'DONOGHUE, Eileen
TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
FILE REFERENCE: DIVER1370-7
CURRENT APPLICATION NUMBER: US/09/866,379
PRIOR APPLICATION NUMBER: US 09/580,515
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 09/318,528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: US 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 09/259,214
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: US 08/910,798
PRIOR FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.1

SEQ ID NO 2
LENGTH: 440
TYPE: PRT
ORGANISM: Escherichia coli
US-09-866-379-2

Query Match 100.0%; Score 2302; DB 2; Length 440;
Best Local Similarity 100.0%; Pred. No. 3.9e-234;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MKAIIIPFLSLIPLTPPSAFASPELKLSEVIVSRHGVRAPRTAKTQMDVTPDAMP 60
DB 1 MKAIIIPFLSLIPLTPPSAFASPELKLSEVIVSRHGVRAPRTAKTQMDVTPDAMP 60
QY 61 TWPVGLGWLTPRGELIAYLGHYQORLVADGLAKKCCPOSGQVAIIADVDETRKTGE 120
DB 61 TWPVGLGWLTPRGELIAYLGHYQORLVADGLAKKCCPOSGQVAIIADVDETRKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPDKTGVCQLDNANVTDAIISRAGSIADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPDKTGVCQLDNANVTDAIISRAGSIADFTGH 180
QY 181 ROTARELERVLPQSNMCLKREKODESCSLTQALPSBLKVSADNVSLTGAVSLASMLT 240
DB 181 ROTARELERVLPQSNMCLKREKODESCSLTQALPSBLKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLQQAQGMPEPMGRITDISHQWNTLLSLHNAQFYLLQRTPEVARSRATPLDLINMA 300
DB 241 EIFLLQQAQGMPEPMGRITDISHQWNTLLSLHNAQFYLLQRTPEVARSRATPLDLINMA 300
QY 301 LTPHPPOQAQVGTLPSTVLPFIAGHDITMLANIGALBELNMTLPQGPDMTPPGSELVFERW 360
DB 301 LTPHPPOQAQVGTLPSTVLPFIAGHDITMLANIGALBELNMTLPQGPDMTPPGSELVFERW 360
QY 361 RLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPPEVGLTLAGCERNAGCISLAGFTQ 420
DB 361 RLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPPEVGLTLAGCERNAGCISLAGFTQ 420
QY 421 IVNEARIPACSLRSHHHHH 440
DB 421 IVNEARIPACSLRSHHHHH 440
```

RESULT 6
US-09-866-379-8
Sequence 8, Application US/09866379
Patent No. 6855365

GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: SHORT, Jay
APPLICANT: KRETZ, Keith
APPLICANT: GRAY, Kevin
APPLICANT: BARTON, Nelson
APPLICANT: GARRETT, James
APPLICANT: O'DONOGHUE, Eileen
TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
FILE REFERENCE: DIVER1370-7
CURRENT APPLICATION NUMBER: US/09/866,379
PRIOR APPLICATION NUMBER: US 09/580,515
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 09/318,528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: US 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 09/259,214
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: US 08/910,798
PRIOR FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.1
LENGTH: 432

TYPE: PRT
ORGANISM: Escherichia coli
US-09-866-379-8

Query Match 97.1%; Score 2235; DB 2; Length 432;
Best Local Similarity 99.5%; Pred. No. 4,56-227;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 MKAILIPFLSLILPLTPQSAFAQSEPELKLSESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
DB 1 MKAILIPFLSLILPLTPQSAFAQSEPELKLSESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
QY 61 TWPVKLGMLTPRGSELIAVIGHYQORLIVADGLAKKGCPOSQGVAIIVADVDETRKTGE 120
DB 61 TWPVKLGMLTPRGSELIAVIGHYQORLIVADGLAKKGCPOSQGVAIIVADVDETRKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTVGCQLDNNAVTDALISRAAGSIADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTVGCQLDNNAVTDALISRAAGSIADFTGH 180
QY 181 RQTAFRELEERVLPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 RQTAFRELEERVLPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLOQAQGMPEBGMKRTIDSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPILDLMMA 300
DB 241 EIFLLOQAQGMPEBGMKRTIDSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPILDLMMA 300
QY 301 LTPHPQKQAYGVTLPTSVLFIAGHDNTLANLGALBLNMTLPQDPDNTPPGSELVPERW 360
DB 301 LTPHPQKQAYGVTLPTSVLFIAGHDNTLANLGALBLNMTLPQDPDNTPPGSELVPERW 360
QY 361 RRLSDNSQWTOVSLVFQTLQOMRDKTPLSINTPPGEVKLTLAGCEERNAQMCSLAGFTQ 420
DB 361 RRLSDNSQWTOVSLVFQTLQOMRDKTPLSINTPPGEVKLTLAGCEERNAQMCSLAGFTQ 420
QY 421 IVNEARIIPACSL 432
DB 421 IVNEARIIPACSL 432
```

RESULT 7
US-09-715-477-1
Sequence 1, Application US/09715477
Patent No. 6841370
GENERAL INFORMATION:
APPLICANT: Lei, Xingen
TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYTASE
FILE REFERENCE: 19603/4031
CURRENT APPLICATION NUMBER: US/09/715,477
CURRENT FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/166,179
PRIOR FILING DATE: 1999-11-18
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 432
TYPE: PRT
ORGANISM: Escherichia coli
US-09-715-477-1

Query Match 96.9%; Score 2230; DB 2; Length 432;
Best Local Similarity 99.3%; Pred. No. 1,56-226;
Matches 429; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 MKAILIPFLSLILPLTPQSAFAQSEPELKLSESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
DB 1 MKAILIPFLSLILPLTPQSAFAQSEPELKLSESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
QY 61 TWPVKLGMLTPRGSELIAVIGHYQORLIVADGLAKKGCPOSQGVAIIVADVDETRKTGE 120
DB 61 TWPVKLGMLTPRGSELIAVIGHYQORLIVADGLAKKGCPOSQGVAIIVADVDETRKTGE 120
```

```
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTVGCQLDNNAVTDALISRAAGSIADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTVGCQLDNNAVTDALISRAAGSIADFTGH 180
QY 181 RQTAFRELEERVLPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 RQTAFRELEERVLPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLOQAQGMPEBGMKRTIDSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPILDLMMA 300
DB 241 EIFLLOQAQGMPEBGMKRTIDSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPILDLMMA 300
QY 301 LTPHPQKQAYGVTLPTSVLFIAGHDNTLANLGALBLNMTLPQDPDNTPPGSELVPERW 360
DB 301 LTPHPQKQAYGVTLPTSVLFIAGHDNTLANLGALBLNMTLPQDPDNTPPGSELVPERW 360
QY 361 RRLSDNSQWTOVSLVFQTLQOMRDKTPLSINTPPGEVKLTLAGCEERNAQMCSLAGFTQ 420
DB 361 RRLSDNSQWTOVSLVFQTLQOMRDKTPLSINTPPGEVKLTLAGCEERNAQMCSLAGFTQ 420
QY 421 IVNEARIIPACSL 432
DB 421 IVNEARIIPACSL 432
```

RESULT 8
US-10-266-041A-1
Sequence 1, Application US/10266041A
Patent No. 6974690
GENERAL INFORMATION:
APPLICANT: Lei, Xingen
TITLE OF INVENTION: PHOSPHATASES WITH IMPROVED PHYTASE ACTIVITY
FILE REFERENCE: 19603/2798
CURRENT APPLICATION NUMBER: US/10/266,041A
CURRENT FILING DATE: 2002-10-07
PRIOR APPLICATION NUMBER: 60/127,032
PRIOR FILING DATE: 1999-03-31
PRIOR APPLICATION NUMBER: 09/540,149
PRIOR FILING DATE: 2000-03-31
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 432
TYPE: PRT
ORGANISM: Escherichia coli
US-10-266-041A-1

Query Match 96.5%; Score 2221; DB 2; Length 432;
Best Local Similarity 99.1%; Pred. No. 1,46-225;
Matches 428; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
QY 1 MKAILIPFLSLILPLTPQSAFAQSEPELKLSESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
DB 1 MKAILIPFLSLILPLTPQSAFAQSEPELKLSESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
QY 61 TWPVKLGMLTPRGSELIAVIGHYQORLIVADGLAKKGCPOSQGVAIIVADVDETRKTGE 120
DB 61 TWPVKLGMLTPRGSELIAVIGHYQORLIVADGLAKKGCPOSQGVAIIVADVDETRKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTVGCQLDNNAVTDALISRAAGSIADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTVGCQLDNNAVTDALISRAAGSIADFTGH 180
QY 181 RQTAFRELEERVLPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 RQTAFRELEERVLPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLOQAQGMPEBGMKRTIDSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPILDLMMA 300
DB 241 EIFLLOQAQGMPEBGMKRTIDSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPILDLMMA 300
QY 301 LTPHPQKQAYGVTLPTSVLFIAGHDNTLANLGALBLNMTLPQDPDNTPPGSELVPERW 360
DB 301 LTPHPQKQAYGVTLPTSVLFIAGHDNTLANLGALBLNMTLPQDPDNTPPGSELVPERW 360
```

Db 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANIAGALELNTLPGQDPNTPPGEGELVFERW 360
QY 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEGVLTLAGCERNAQGCISLAGFTQ 420
Db 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEGVLTLAGCERNAQGCISLAGFTQ 420
QY 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432

RESULT 9
US-09-540-149A-1
; Sequence 1, Application US/09540149A
; Patent No. 6511699
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: ENZYMES WITH IMPROVED PHYTASE ACTIVITY
; FILE REFERENCE: 19603/2791
; CURRENT APPLICATION NUMBER: US/09/540,149A
; CURRENT FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 60/127,032
; PRIOR FILING DATE: 1998-03-31
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (433)
; OTHER INFORMATION: Xaa at position 433 in this sequence is unknown
US-09-540-149A-1

Query Match 96.5%; Score 2221; DB 2; Length 433;
Best Local Similarity 99.1%; Pred. No. 1,4e-225;
Matches 428; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MKAILIPLSLILPITPQSAFAQSEPELKESVIVSRHGVAAPTKATQIMQDVTDPDAMP 60
Db 1 MKAILIPLSLILPITPQSAFAQSEPELKESVIVSRHGVAAPTKATQIMQDVTDPDAMP 60
QY 61 TWPVKGWLTTRGSELIVLYGHYQRLVADGLAKKGCPOGQVAIADVDERTKTGE 120
Db 61 TWPVKGWLTTRGSELIVLYGHYQRLVADGLAKKGCPOGQVAIADVDERTKTGE 120
QY 121 AFAAGLADCAITVHTQADTSSPDLFNPPLKTGVQOLNANVTDAIISRAGSIAIDFTGH 180
Db 121 AFAAGLADCAITVHTQADTSSPDLFNPPLKTGVQOLNANVTDAIISRAGSIAIDFTGH 180
QY 121 AFAAGLADCAITVHTQADTSSPDLFNPPLKTGVQOLNANVTDAIISRAGSIAIDFTGH 180
Db 121 AFAAGLADCAITVHTQADTSSPDLFNPPLKTGVQOLNANVTDAIISRAGSIAIDFTGH 180
QY 181 RQTAARELERLVNFPQSNLCKREKQDSCSIQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 RQTAARELERLVNFPQSNLCKREKQDSCSIQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLQQAQGMPEBPGWGRITDISHQNTLLSIHNAQFYLLQRTPEVARSRATPDLIDLIMAA 300
Db 241 EIFLLQQAQGMPEBPGWGRITDISHQNTLLSIHNAQFYLLQRTPEVARSRATPDLIDLIMAA 300
QY 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANIAGALELNTLPGQDPNTPPGEGELVFERW 360
Db 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANIAGALELNTLPGQDPNTPPGEGELVFERW 360
QY 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEGVLTLAGCERNAQGCISLAGFTQ 420
Db 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEGVLTLAGCERNAQGCISLAGFTQ 420
QY 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432

RESULT 10

US-09-715-477-3
; Sequence 3, Application US/09715477
; Patent No. 6841370
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYTASE
; FILE REFERENCE: 19603/4031
; CURRENT APPLICATION NUMBER: US/09/715,477
; CURRENT FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,179
; PRIOR FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-715-477-3

Query Match 96.0%; Score 2210; DB 2; Length 432;
Best Local Similarity 98.6%; Pred. No. 2e-224;
Matches 426; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MKAILIPLSLILPITPQSAFAQSEPELKESVIVSRHGVAAPTKATQIMQDVTDPDAMP 60
Db 1 MKAILIPLSLILPITPQSAFAQSEPELKESVIVSRHGVAAPTKATQIMQDVTDPDAMP 60
QY 61 TWPVKGWLTTRGSELIVLYGHYQRLVADGLAKKGCPOGQVAIADVDERTKTGE 120
Db 61 TWPVKGWLTTRGSELIVLYGHYQRLVADGLAKKGCPOGQVAIADVDERTKTGE 120
QY 121 AFAAGLADCAITVHTQADTSSPDLFNPPLKTGVQOLNANVTDAIISRAGSIAIDFTGH 180
Db 121 AFAAGLADCAITVHTQADTSSPDLFNPPLKTGVQOLNANVTDAIISRAGSIAIDFTGH 180
QY 181 RQTAARELERLVNFPQSNLCKREKQDSCSIQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 RQTAARELERLVNFPQSNLCKREKQDSCSIQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLQQAQGMPEBPGWGRITDISHQNTLLSIHNAQFYLLQRTPEVARSRATPDLIDLIMAA 300
Db 241 EIFLLQQAQGMPEBPGWGRITDISHQNTLLSIHNAQFYLLQRTPEVARSRATPDLIDLIMAA 300
QY 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANIAGALELNTLPGQDPNTPPGEGELVFERW 360
Db 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANIAGALELNTLPGQDPNTPPGEGELVFERW 360
QY 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEGVLTLAGCERNAQGCISLAGFTQ 420
Db 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEGVLTLAGCERNAQGCISLAGFTQ 420
QY 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432

RESULT 11
US-08-910-798-2
; Sequence 2, Application US/08910798
; Patent No. 5876997
; GENERAL INFORMATION:
; APPLICANT: KRETZ
; TITLE OF INVENTION: NOVEL PHYTASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/910,798
FILING DATE: August 13, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: HALL, PH.D., LISA A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 09010/029001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-910-798-2

Query Match 95.3%; Score 2194.5; DB 1; Length 423;
Best Local Similarity 96.1%; Pred. No. 8.3e-223;
Matches 423; Conservative 0; Mismatches 0; Indels 17; Gaps 1;

QY 1 MKALLIPFLSLILPLTPQSAFAOSEPELKESVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
DB 1 MKALLIPFLSLILPLTPQSAFAOSEPELKESVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
QY 61 TWPKLGLTPRGSELIAVIGHYQORLVADGLLAKKGCPOSGOVAIADVDETRKTGE 120
DB 61 TWPKLGLTPRGSELIAVIGHYQORLVADGLLAKKGCPOSGOVAIADVDETRKTGE 119
QY 121 AFAAGLAPDCAIVHTQADTSSPDLFNPPLKTVGCQLDNANVTDAIISRAGGS1ADFTGH 180
DB 120 -----QADTSSPDLFNPPLKTVGCQLDNANVTDAIISRAGGS1ADFTGH 163
QY 181 RQTFRELERLVNPPQSNLCIKREKODSCSLTQALSELKVSADNVSLTGAVALSLMLT 240
DB 164 RQTFRELERLVNPPQSNLCIKREKODSCSLTQALSELKVSADNVSLTGAVALSLMLT 223
QY 241 EIFLLQQAQGMPEBQWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
DB 224 EIFLLQQAQGMPEBQWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIMAA 283
QY 301 LTPHPQKQAVGVTLPTSVLFIAGHDNLANLGGALBLNMTLPGQPDNTPPGGELVFERW 360
DB 284 LTPHPQKQAVGVTLPTSVLFIAGHDNLANLGGALBLNMTLPGQPDNTPPGGELVFERW 343
QY 361 RRLSDNSQWIOVSLVFOQLQMRDXTPLSLNTPPGEVKLTLAGGEERNAQMCSLAGFTQ 420
DB 344 RRLSDNSQWIOVSLVFOQLQMRDXTPLSLNTPPGEVKLTLAGGEERNAQMCSLAGFTQ 403
QY 421 IVNEARIPACSLRSHHHHH 440
DB 404 IVNEARIPACSLRSHHHHH 423

RESULT 12
US-09-866-379-10
Sequence 10, Application US/09866379
Patent No. 6855365
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: SHORT, Jay
APPLICANT: KRETZ, Keith
APPLICANT: GRAY, Kevin
APPLICANT: BARTON, Nelson
APPLICANT: GARRETT, James
APPLICANT: O'DONOGHUE, Eileen
TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
FILE REFERENCE: DIVER1370-7

CURRENT APPLICATION NUMBER: US/09/866,379
CURRENT FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: US 09/580,515
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 09/318,528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: US 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 09/259,214
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: US 08/910,798
PRIOR FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.1
SEQ ID NO 10
LENGTH: 430
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Modified phytase
US-09-866-379-10

Query Match 93.0%; Score 2141; DB 2; Length 430;
Best Local Similarity 97.2%; Pred. No. 3.9e-217;
Matches 420; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY 1 MKALLIPFLSLILPLTPQSAFAOSEPELKESVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
DB 1 MKALLIPFLSLILPLTPQSAFAOSEPELKESVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
QY 61 TWPKLGLTPRGSELIAVIGHYQORLVADGLLAKKGCPOSGOVAIADVDETRKTGE 120
DB 61 TWPKLGLTPRGSELIAVIGHYQORLVADGLLAKKGCPOSGOVAIADVDETRKTGE 120
QY 121 AFAAGLAPDCAIVHTQADTSSPDLFNPPLKTVGCQLDNANVTDAIISRAGGS1ADFTGH 180
DB 121 AFAAGLAPDCAIVHTQADTSSPDLFNPPLKTVGCQLDNANVTDAIISRAGGS1ADFTGH 179
QY 181 RQTFRELERLVNPPQSNLCIKREKODSCSLTQALSELKVSADNVSLTGAVALSLMLT 240
DB 180 RQTFRELERLVNPPQSNLCIKREKODSCSLTQALSELKVSADNVSLTGAVALSLMLT 239
QY 241 EIFLLQQAQGMPEBQWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
DB 240 EIFLLQQAQGMPEBQWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIMAA 298
QY 301 LTPHPQKQAVGVTLPTSVLFIAGHDNLANLGGALBLNMTLPGQPDNTPPGGELVFERW 360
DB 299 LTPHPQKQAVGVTLPTSVLFIAGHDNLANLGGALBLNMTLPGQPDNTPPGGELVFERW 358
QY 361 RRLSDNSQWIOVSLVFOQLQMRDXTPLSLNTPPGEVKLTLAGGEERNAQMCSLAGFTQ 420
DB 359 RRLSDNSQWIOVSLVFOQLQMRDXTPLSLNTPPGEVKLTLAGGEERNAQMCSLAGFTQ 418
QY 421 IVNEARIPACSL 432
DB 419 IVNEARIPACSL 430

RESULT 13
US-09-489-039A-7512
Sequence 7512, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342

SEQ ID NO 7512
LENGTH: 421
TYPE: PR
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7512

Query Match 22.9%; Score 528; DB 2; Length 421;
Best Local Similarity 32.1%; Pred. No. 1.1e-46;
Matches 134; Conservative 69; Mismatches 184; Indels 30; Gaps 10;

QY 19 SAFAOSE---PE-LKTESVIVSRHGVRAP-TKATQMODVTPDAMPWPKLGMLTR 72
DB 23 SAGAAQAKAEQEGYQOQVLLMSRHNLRAPLANNGLVLEGSTAXAMPQMDVPGQLTKK 82
QY 73 GSELIAVGHYORQLVADGLAKKGCPSQGVALLADVDETRTKTGEPAGLAPDCAI 132
DB 83 GGVLVVGHYHREMLAQOKLVTSGECPENAVVAVNSLQRTVATAQPFITGAPPGCGI 142
QY 133 TVHTQADTSPDPLFNPPLKTVGCOLDNNAVTDAL-LSRAGSIADEPTGHRQTAPELEEV 191
DB 143 PVHHQPMQMDPTNPNVITTDSPAPREKALQAMEKRGQMOLTE-----SYKLEETM 195
QY 192 LNFPSQNLCLKREKQDESCSLQALPSBLKVSADNVSLTGAVALSLMTEIFLLQQAQGM 251
DB 196 IDYRNSPSC---KEKVCSLSEKQOTFSAGYQGERGVSGPLKVGNSLVDAFTLLGYTSGF 251
QY 252 P--EFGMRITDSHQMNTLSLHNAQFYLQRTPEVASRATPLDLIMALTTPHPQKQ 309
DB 252 PDDQVAMGBIASDKQMRVLSKLKNGYODSLFTSVAVNAQVAPLVXYIDNALVGBGASK- 310
QY 310 AVGVTLPTSVLFIAGHDITLANLGGALBELN-WTLPGQPNTPGGEIVPEMRRLSDNSQ 368
DB 311 -----AKVTLVHDSNISLTLALDCKPYQLPGQYERTITIGGLTLFQRHDSAGNRD 363
QY 369 WIVQSLVEFQIQMDKTPSLNTPPGEVKLTLAGCEERNAGQMSLAGFTQIVNEA 425
DB 364 LMKIEVYVQSTEQALNADALTLOAPFQRTVTLANGCPV-DDGGFPLETFKVVINEA 419

RESULT 14
US-09-489-039A-13501
Sequence 13501, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:

APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 13501
LENGTH: 522
TYPE: PR
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13501

Query Match 16.3%; Score 376; DB 2; Length 522;
Best Local Similarity 25.8%; Pred. No. 1.1e-30;
Matches 115; Conservative 75; Mismatches 209; Indels 46; Gaps 15;

QY 2 KAILPLSLIPLTP-OSAPFAQSEPELKESVIVSRHGVRAPTKAT-QLMQDVTTPDAW 59
DB 107 QGLRLFLFACALPLALQSAAL-ADMQLKVELSRHGIRPPTAGNRBAIIBATGKRPW 163
QY 60 PTWPKVGLTPRGGELIAVGHYORQLVADGLAKKGCPSQGVALLADVDETRTKTG 119
DB 164 TEWTHDELTGHGAAVAVNKGRAEGQHYROLGLL-QAGCPTABSIIVYRASPLOQTRATA 222
QY 120 EAPFAGLAPDCAITVHTQADTSPDPLFNPPLKTVGCOLDNNAVTDALISRAGSIADEPTG 179

DB 223 QALVNDGAPGCGVAIHY--SGDADPLFQTDKFAATQTDPARQLAAVKEKAG----DLAQ 276
QY 180 HQQTAFRELERVLANFPQSNLCLKREKQDESCSLQALPSBLKVSADNVSLTGAVALSLM 239
DB 277 RROA-----LAPTIQLKQAVC---QADKPCPIPDT-PWQVQSKSGKTTISGLVMAVM 327
QY 240 TEIFLLQQAQGMPE--EPGMRITDSHQMNTLSLHNAQFYLQRTPEVASRATPLDLTI 297
DB 328 VETLRLGSENLPLSQLAMGKTQARQTTALPLPLETNEYDLTNDVLYTRQKRGSVLLNAM 387
QY 298 MAALTTPHPPOKQAVGVTLPTSVLFIAGHDITLANLGGALBELNLTLPQ-QPDNTPGGEIV 356
DB 388 LDGVPEANPNVRW-----LLVADITNIAVRLTMINSSWQLPGVSRGNIIPGGSILV 439
QY 357 FERRRLSDNSQMIQVSLVFQTLQMDKTPSLNTPPGEVKLT-----LAGCEERNAG 411
DB 440 LERWENAKSGERYLVRYQAOGLDLR-----RLQTPDAQHMLRQEWHPGCRQTDVGT 494
QY 412 MC---SLAGFTQIVNEARIPACSL 432
DB 495 LCPFOAITLALGQRIIDRSSAPAVAM 519

RESULT 15
US-09-044-718-3
Sequence 3, Application US/09044718
Patent No. 6391605
GENERAL INFORMATION:

APPLICANT: KOSTREWA, Dirk
APPLICANT: PASAMONTES, Luis
APPLICANT: TOMSCHY, Andrea
APPLICANT: VAN LOON, Adolphus
APPLICANT: VOSEL, Kurt
APPLICANT: WYSS, Markus
TITLE OF INVENTION: MODIFIED PHYTASES
FILE REFERENCE: Modified Phytases
CURRENT APPLICATION NUMBER: US/09/044,718
CURRENT FILING DATE: 1998-03-19
PRIOR APPLICATION NUMBER: EP 97810175.6
PRIOR FILING DATE: 1997-03-25
NUMBER OF SEQ ID NOS: 82
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 439
TYPE: PR
ORGANISM: Aspergillus fumigatus
US-09-044-718-3

Query Match 5.2%; Score 119; DB 2; Length 439;
Best Local Similarity 21.7%; Pred. No. 0.002;
Matches 102; Conservative 70; Mismatches 183; Indels 114; Gaps 24;

QY 7 PFLSLIPLTPQSAFAQSEPELKESVIVSRHGVRAP-----KATQMODVTPDA--- 58
DB 26 PFLSLEDELSSVSLPK---DCKITVQVLSRHGARPTSSKSKYKVLVAIQANATDF 82
QY 59 -----WETWPKLGM--LTPRGGELIAYIG--HYORQLVADGLAKKGCPSQGVAIL 108
DB 83 KGKFAFLKTYVTLGADDLTFPGEQQLVNSGIFKFORKALARSVP-----FIR 132
QY 109 ADVBERIRKTEAARAGL-----ADCAITVHTQADTSPDPLFNPPLKTVG 154
DB 133 ASGSDRYIASGEKTEIEFQAKLADPGATYNNRAAIAISYII-PBSRTFV-----ATLHGV 186
QY 155 COLDNNAVTDALISRAGSI-ADPTGHRQTAPELEERLANFPQSNLCLKREKQDESCSLT 213
DB 187 C-----TKRAQQLDDEVANFT-----ALRAP--DIRAKELHPGVTLT 225
QY 214 QALPSEL--KVSADNVSLTGAVALSLMTEIFLLQQAQGMPEPGWGRITDSHQMNTLSL 271
DB 226 DEDVAVSLMDKSPFTVARTSDASQLSPFCQLFTNBE-----KKYNYLQSL 271
QY 272 HNAQFYLQRTPEVASRATPLDLIMALTTPHPQK-----AVGVTLPTSVLPIA 323

Db 272 --GKYYGVGAGNPIGPAQIGFTNEILARLTRSPVQDHTSTNSTVSNPATFPPLNATMYV 329
QY 324 --GHDITLANTGALIELNMTLPGQPDNTPPGGELVFERWRRLSD-NSQWT--QVSLVEQ 377
Db 330 DFSDNSMWSIFPAL-----GLYNGTEPLSRTSVESAKELDGYASWVVPFGARAYFE 382
QY 378 TLQGMRDKTPLSLNTPPGEVKLTLAGCERNAGWCSLAGFTQIVNEAR 426
Db 383 TMQCKSEKEPL-VRALINDRVVPLHGCV-DKLGRCCKLNDPVGKLSWAR 429

Search completed: June 13, 2006, 10:30:45
Job time : 38.3303 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 14, 2006, 11:48:02 ; Search time 4986.33 Seconds
(without alignments) 7401.587 Million cell updates/sec

Title: US-10-601-319-2

Perfect score: 2302

Sequence: 1 MKALLIFPLSLILPLTPSA.....IVNEARIPACSLRHHHHH 440

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 48236798 seqs, 2795965780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xip
-Q=/abs/ABSSWB.epool/US10601319/runat.12062006.180052.27516/app.query.fasta_1
-DB=EST -QFMT=fastdb -SUFFIX=est -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biom62 -TRANS=human4.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs03p
-USER=US10601319 @CGN_1_12067 @runat.12062006.180052.27516 -NCPU=6 -ICPU=3
-NO_WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: gb_est1:
2: gb_est3:
3: gb_est4:
4: gb_est5:
5: gb_est6:
6: gb_est7:
7: gb_est8:
8: gb_est9:
9: gb_est10:
10: gb_est11:
11: gb_est12:
12: gb_est13:
13: gb_est14:
14: gb_est15:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1044	45.4	616	2	BU074127 BU074127
2	1036	45.0	853	13	CU662734 PRI0142B
3	913	39.7	529	3	BU113770 SJABUG02
4	896	38.9	746	4	CA093060 SCCCL200

C	5	846	36.8	595	5	CP326092	CP326092 JMT1--05-
C	6	838	36.4	895	8	CN762997	CN762997 ID0AA5DC
C	7	767	33.3	868	8	CN754382	CN754382 ID0AA12D
C	8	757	32.9	864	8	CN757565	CN757565 ID0AA1DF
C	9	643	27.9	383	7	AM036132	AM036132 EST274508
C	10	643	27.9	383	7	AM036134	AM036134 EST274510
C	11	641	27.8	393	7	BE520240	BE520240 M11B12STM
C	12	591.5	22.7	531	3	BU113769	BU113769 SJABUG01
C	13	522	22.7	706	8	CN759004	CN759004 ID0AA24B
C	14	511.5	22.2	354	1	AA545747	AA545747 HBMSF1B4-
C	15	442	19.2	707	2	BU619443	BU619443 BJ619443
C	16	392	17.0	320	7	BE366403	BE366403 EST40781
C	17	374.5	16.3	1175	9	CP943192	CP943192 LamiGeGeG
C	18	355	15.4	208	2	BM409016	BM409016 EST183343
C	19	355	15.4	208	2	BM412806	BM412806 EST587133
C	20	350.5	15.2	821	13	CL665831	CL665831 PRI014C_E
C	21	345.5	15.0	842	13	CL665745	CL665745 PRI014C_E
C	22	339	14.7	546	2	BU040796	BU040796 BJ040796
C	23	297	12.6	757	4	CA093061	CA093061 SCCCL200
C	24	287	12.5	657	2	BG457132	BG457132 NF100C05P
C	25	284	12.3	1171	8	CV672454	CV672454 RETVSU_01
C	26	278	12.1	179	7	BE461872	BE461872 EST413251
C	27	269	11.7	508	2	BU030203	BU030203 BJ030203
C	28	228.5	9.9	480	1	AL586973	AL586973 AL586973
C	29	225	9.8	254	7	BE520241	BE520241 M11B12XTM
C	30	225	9.8	470	9	DN201839	DN201839 USDA-FP_1
C	31	219.5	9.5	692	5	CK564106	CK564106 rswpD0_00
C	32	215	9.3	772	14	DX010757	DX010757 KBXB001C0
C	33	195	8.5	551	5	CK535965	CK535965 rswpD0_00
C	34	186	8.1	692	9	CK700688	CK700688 Hm_L6w_04
C	35	185.5	8.1	575	5	CK543282	CK543282 rswpD0_01
C	36	183	7.9	687	10	DV079873	DV079873 I13Fm1dg
C	37	172.5	7.5	696	1	AU004475	AU004475 AU004475
C	38	172	7.5	734	5	CK507686	CK507686 rswd0_00
C	39	163	7.1	769	8	CO386885	CO386885 AGENCOUTT
C	40	153	6.6	863	8	CV672453	CV672453 RETVSJ_01
C	41	150	6.5	1044	6	CNS080M3	BN211319 Single_7e
C	42	144.5	6.3	645	2	BU037937	BU037937 BU037937
C	43	139.5	6.1	1430	6	CR615614	CR615614 full_1eng
C	44	137.5	6.0	2201	6	CR926080	CR926080 Pongo Pyg
C	45	133	5.8	1125	10	DM035689	DM035689 CFM261-G0

ALIGNMENTS

RESULT 1
BU074127/c 616 bp mRNA linear EST 29-SEP-2003
LOCUS
DEFINITION
BU074127 NIBB Mochii normalized Xenopus tailbud library Xenopus
laevis cDNA clone X1090506 5', mRNA sequence.

ACCESSION
BU074127
VERSION
BU074127.1 GI:17504316
KEYWORDS
SOURCE
ORGANISM
Xenopus laevis (African clawed frog)

REFERENCE
1 (bases 1 to 616)
Kikayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-I, T. and
Kohara, Y.
Expressed genes in X. laevis embryo
Unpublished (2001)
Contact: Tadasi Shin-I
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855

TITLE
JOURNAL
COMMENT

Email: tshin@genie.nig.ac.jp
The information of this clone is available through the following
URL:
http://xenopus.nibb.ac.jp.

FEATURES
source

Location/Qualifiers
1. .616
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="XL090106"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/clone_lib="NIBB Mochizuki normalized Xenopus tailbud library"

ORIGIN

Alignment Scores:
Pred. No.: 5,316-96 Length: 616
Score: 1044.00 Matches: 204
Percent Similarity: 99.5% Conservative: 0
Best Local Similarity: 99.5% Mismatches: 1
Query Match: 45.4% Indels: 0
DB: 2 Gaps: 0

US-10-601-319-2 (1-440) x B0704127 (1-616)

QY 89 ValAlaApGgLyLeuLeuAlaAlaYsGgLyCyApProGInserGlyGlnValAlaIleIle 108
DB 615 GTAGCCGACGGATTGCTGGCGAAGAGGCTGCCCGACGTCTGGTCAAGTCGCCATTATT 556
QY 109 AlAAspValaApGgLyGlyThAArgLyThArgLyGlyAlaApeAlaAlaGlyLeuAlaPro 128
DB 555 GCTGATGTCGACGCGGACCGGTAAACAGCGCAAGCCTTCGCCCGCGGCTGGCACT 496
QY 129 AspCyAlaIleIleThrValaIleThrGlnAlaApThrSerSerProApProLeuPheAsn 148
DB 495 GACTGTGCAATACCGTACATACCGACATACCTCACTCCGATCCGATTATTAT 436
QY 149 ProLeuYsThrGlyValaCyAGlnLeuAspAsnAlaAsnValaIleThrAspAlaIleLeuSer 168
DB 435 CCTCTAAABAACTGGGGTTGGCAACTGTGATACCGCAAGTGAAGTCAAGCCATCCTCAGC 376
QY 169 ArgAlaGlyGlySerIleAlaAspPheThrGlyIleAArgGlnThrAlaPheArgGlyLeu 188
DB 375 AGGCGACGAGGGTCAATTGCTGACTTACCGGCAACCGCGCTTTCGCGAACTG 316
QY 189 GlnUArgValaLeuAspPheProGlnSerAsnLeuCySLeuYsArgGlnUlyGlnhAspGln 208
DB 315 GAACGGGCTTAATTTTCCGCAATCAACTTGTCTTAAACGTGAAGAACGAGACAA 256
QY 209 SerCySerLeuThrGlnAlaLeuProSerGlnLeuYsValaSerAlaAspAsnValaSer 228
DB 255 AGCTGTTCAATTAAACGACGATTAACATCGGAACCAAGTGAAGCGCCGACAAATGTCTCA 196
QY 229 LeuThrGlyValaValaSerLeuAlaSerMetLeuThrGlnIlePheLeuLeuGlnAla 248
DB 195 TTAAACCGGTGGGTAAGCTCGCATCAATGCTGACGAGGATATTCTCTCGCAACAGCA 136
QY 249 GlnGlyMetProGlnUProGlyYTrpGlyYArgIleThrAspSerHisGlnIleThrAsnThrLeu 268
DB 135 CTGGGAAGTCCGAGACCGGGGTGGGAAGATCAACCAATTCACACCAATGGAAACCTTG 76
QY 269 LeuSerLeuHisAsnAlaGlnPheThrLeuLeuGlnArgThrProGlnValaAlaArgSer 288
DB 75 CTAAAGTTTGCAATACGCGCAATTTATTGTGACAAACGACGCGAGAGTGGCCGCGAGC 16
QY 289 ArgAlaThrProLeu 293
DB 15 CGCGCCACCCCGTTA 1

RESULT 2
LOCUS CL662734 853 bp DNA linear GSS 09-JUL-2004
DEFINITION PR10142b.B21 (853) Mixed stage fosmid library of P. pacificus var. California Pristionchus pacificus genomic survey sequence.
ACCESSION CL662734

VERSION : CL662734.1 GI:50150877
KEYWORDS : GSS.
SOURCE : Pristionchus pacificus
ORGANISM : Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Neodiplogasteridae; Pristionchus.
REFERENCE : 1 (bases 1 to 853)
AUTHORS : Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
TITLE : Appads: an Acedb database for the nematode satellite organism Pristionchus pacificus
JOURNAL : Nucleic Acids Res. 32 (1), D421-D422 (2004)
PUBMED : 14681447
COMMENT : Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel.: 00497071601378
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
This library was generated at Caltech, Pasadena, USA and end sequenced at Vancouver, Canada.
Seq primer: T7
Class: fosmid ends.

FEATURES
source

Location/Qualifiers
1. .853
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Mixed stage fosmid library of P. pacificus var. California"
/note="Vector: pBf105-5 Fosmid vector"

ORIGIN

Alignment Scores:
Pred. No.: 5,736-95 Length: 853
Score: 1036.00 Matches: 196
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 45.0% Indels: 0
DB: 13 Gaps: 0

US-10-601-319-2 (1-440) x CL662734 (1-853)

QY 237 SerMetLeuThrGlnIlePheLeuLeuGlnAlaGlnGlyMetProGlnUProGlyYTrp 256
DB 2 TCATAGTGAACGGAGATATTCTCTGCAACACACAGGAAATCCGAGCCGGGTGG 61
QY 257 GlnYArgIleThrAspSerHisGlnIleThrAsnThrLeuLeuSerLeuHisAsnAlaGlnPhe 276
DB 62 GGAAGATACCGGATTCAACACAGTGAACACTTGTCTAAAGTTGCATTAACGCCAATTT 121
QY 277 TyrLeuLeuGlnArgThrProGlnValaAlaArgSerArgAlaThrProLeuLeuAspLeu 296
DB 122 TATTGTGACAAACCGACCGCAGAGTTGCCGACGCGCCGACCCGCTTATTGATTG 181
QY 297 IleMetAlaIleLeuThrProHisProProGlnIlyGlnAlaYArgIleValaIleThrLeuPro 316
DB 182 ATCATGGGAGGTTGAGCCCATTCACCGAAGAAACGGGCTATGGTGTACATTACC 241
QY 317 ThrSerValaLeuPheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyValaLeu 336
DB 242 ACTTCAGTACTGTTATTATGGCGGACAGATCTAATCTGGCAAAATCTCGCGGCGCACTG 301
QY 337 GlnLeuAsnThrPheThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlyLeuVal 356
DB 302 GAGCTCAACTGAGCGCTTCCCGGTACGCGGATTAACGCGCCAGGGTGAACCTGGTG 361
QY 357 PheGlnUArgTTPAAGTGAArgLeuSerAspAsnSerGlnTrpIleGlnValaSerLeuValaPhe 376
DB 362 TTGAACGCTGGCGCTGAGTGAAGATACAGCCAGTGAATTCAGTTCCCTGGTCTTC 421
QY 377 GlnThrLeuGlnGlnMetArgAspLyThrProLeuSerLeuAsnThrProProGlyGln 396

Db 422 CAGACTTTACAGACGATGCGATATAAAGCGCGCTGCTCATTAATACGCGCCGGAGAG 481
 Qy 397 VallyleuThreua1aGlyCyseGluGuaArgaAa1aGlnGlyMetCyseSerleu1a 416
 Db 482 GGAAGACGACCTGCGAGATGTATAAGAGCGAAGAGCCGACGCGCATGTGTGTTGGCA 541
 Qy 417 GlyPheThrGln1aValaAaGlu1aAArg1aPro1aCyseSerleu 432
 Db 542 GGTTTTACCAATCGTGAATGAACGACCATACCGCGCTGCAGTTTG 589
 RESULT 3
 B0713770
 LOCUS
 DEFINITION
 B0713770 529 bp mRNA linear EST 23-OCT-2003
 SJABU02 Adult SJC 7/94 Schistosoma japonicum cDNA similar to
 pbd1JDKP1A chain A, Crystal Structure Of Phytase Complex Of
 Bacteriophage Coli Phytase At Ph 6.6. Phytase Is Bound With Its
 3-Phosphate In The Active Site. Hg2+ Cation Acts As An
 Intermedular Bridge, mRNA sequence.
 B0713770
 B0713770.1 GI:28321126
 EST
 Schistosoma japonicum
 Schistosoma japonicum
 Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
 Strigeidae; Schistosomatidae; Schistosomidae; Schistosoma.
 1 (bases 1 to 529)
 Hu, W., Yan, Q., Shen, D.K., Liu, F., Zhu, Z.D., Song, H.D., Xu, X.R.,
 Wang, Z.J., Kong, Y.P., Zeng, L.C., Wu, J., Zhang, X., Wang, J.J.,
 Xu, X.N., Wang, S.Y., Fu, G., Zhang, X.L., Wang, Z.Q., Birdley, P.J.,
 McMahon, D.P., Xue, C.L., Feng, Z., Chen, Z., and Han, Z.G.
 Evolutionary and biomedical implications of a Schistosoma japonicum
 complementary DNA resource
 Nat. Genet. 35 (2), 139-147 (2003)
 12973349
 JOURNAL
 PUBMED
 COMMENT
 Contact: Zeguang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 201203, P. R. China
 Tel: 86-21-50801919 (ex. 45)
 Fax: 86-21-50801922
 Email: hanzg@chgc.sh.cn.
 Location/Qualifiers
 1..529
 /organism="Schistosoma japonicum"
 /mol_type="mRNA"
 /strain="Chinese (Anhui) strain"
 /db_xref="taxon:6182"
 /sex="Male and female"
 /tissue_type="Whole body"
 /dev_stage="Adult worms"
 /lab_host="Mouse and rabbit"
 /clone_lib="Adult SJC 7/94"
 /note="Vector: Lambda ZAP-II XR.; Site 1: EcoR I; Site 2:
 XhoI I; Several hundred adult Schistosoma japonicum
 (Anhui, P.R. China, strain), of mixed sex, were perfused
 from the mesenteries of experimentally infected mice and
 rabbits at the Queensland Institute of Medical Research,
 Brisbane, Australia (QIMR), and stored for several months
 in liquid nitrogen. Subsequently, mRNA was isolated at the
 QIMR from lysates of these worms by oligo dT
 chromatography, using a kit from Pharmacia. The mRNA was
 then shipped to Clontech, Palo Alto, CA, USA, who
 constructed a cDNA library. First strand synthesis was
 primed with an oligo-dT-XhoI-primer and synthesized using
 M-MuV reverse transcriptase. Second strand synthesis was
 accomplished with RNase H and T4 DNA polymerase. The
 double stranded cDNA was ligated to EcoRI linkers,
 digested with EcoRI and XhoI, and ligated into the
 phagemid vector lambda ZAP II XR. After construction of
 this directional library by Clontech, it was returned to
 the QIMR. During analysis of the library at the QIMR, we
 have found that a small percentage, 2% to 3%, of the

clones contain inserts that appear to be highly homologous
 to sequences from salmonid fishes, as determined by
 homology comparisons using BLAST and by Southern
 hybridization analysis to genomic DNA from salmon (Sigma
 Chemical Co., St. Louis, MO) under stringent washing
 conditions. The remainder of the clones appear to contain
 S. japonicum sequences."

ORIGIN
 Alignment Scores:
 Pred. No.: 1,086-82 Length: 529
 Score: 913.00 Matches: 174
 Percent Similarity: 98.9% Conservative: 0
 Best Local Similarity: 98.9% Mismatches: 2
 Query Match: 39.7% Indels: 0
 DB: 3 Gaps: 0
 US-10-601-319-2 (1-440) x B0713770 (1-529)
 Qy 224 A1aAaPaenValSerleuThrGlyAlaValSerleu1aSerMetleuThrGlu1lePhe 243
 Db 1 GCCGACAAATGCTCATTAACGGTGGTAAAGCTCGCATCAATCTGACGAGATATT 60
 Qy 244 LeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArg1leThrAspSerHis 263
 Db 61 CTCCTGCAACAAGCACAGGAAATGCCGAGCGGGGTGGGAAGATACCGATTCACAC 120
 Qy 264 GlnTPaenThrLeuLeuSerleuHisAaenAlaGlnPheTyrLeuLeuGlnAArgThrPro 283
 Db 121 CAGTGAACACCTTGCTAATGTTGATTAACCGCAATTTATTTCTCAACAGCAGCCCA 180
 Qy 284 GluValAlaAArgSerAArgAlaThrProleuAaPleu1emeAlaAlaLeuThrPro 303
 Db 181 GAGGTGCGCGAGCGCGCCACCCGTTATTGATTCAGACAGCGTTGACGCC 240
 Qy 304 HisProGln1aValGlnAlaTyrGlyAlaThrleuProThrSerValleuPhe1leAla 323
 Db 241 CATCAACGCAAAAACAGCGGTATGTGTGACATTAACCATTCATGCTGTATTTCGCC 300
 Qy 324 GlyHisAaPThrAenLeu1aAaenGlyGlyAlaLeuGluLeuAaenTrpThrleuPro 343
 Db 301 GGAACAGATACTAATCTGCAAAATCTCGCGCGGACATGAGCTCACTGACGCTTCCC 360
 Qy 344 GlyGlnProAaPaenThrProProGlyGlyGluLeuValPheGlnThrleuGlnMetArg 363
 Db 361 GGTACCGCGATTAACACCGCCGAGGTGTGAACGTGTTGAACGCTGGCGTGCCTA 420
 Qy 364 SerAaPaenSerGlnTrp1leGlnValSerleuValPheGlnThrleuGlnMetArg 383
 Db 421 AGCGATTAACAGCGATGATTCAGGTTTCGCTGCTTCACGACTTTACAGCAGATGCGT 480
 Qy 384 Asp1yThrProLeuSerleuAaenThrProProGlyGlyVal1yAeLeu 399
 Db 481 GATTAATAACCGCGCTGTCTAATAATACGCGCCGAGAGGTAACCTG 528
 RESULT 4
 CA093060
 LOCUS
 DEFINITION
 CA093060 746 bp mRNA linear EST 23-SEP-2003
 SCCCL2001D10.b Cl2 Saccharum officinarum cDNA clone SCCCL2001D10
 3', mRNA sequence.
 CA093060
 ACCESSION
 CA093060.1 GI:34946367
 VERSION
 CA093060.1
 KEYWORDS
 EST.
 SOURCE
 Saccharum officinarum
 ORGANISM
 Saccharum officinarum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
 clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
 complex.
 1 (bases 1 to 746)
 da Silva, A.L., da Silva, F.R., Kemper, E.L., and Arruda, P.
 The libraries that made SUCSEST
 Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

QY 310 AlaTYRGLVAlThrleuProThrsSerValleuPhelelaaglyYHAspThrAsnleu 329
| | | | |
Db 475 GGGTATGGTGTGACATTACCACTTAGTGCCTTTATGCGGACACCAATTAATCTG 416
QY 330 AlaAsnleuGLVGLYAlaleuGLuleuAsnTrpThrleuProGLYInProAspAsnThr 349
| | | | |
Db 415 GGAATCTCGGGGGGCACTGGAGCTCAACTGACCCCTCCCGTACGCCGATTAACAG 356
QY 350 ProProGLYGLYGLuleuValPheGLUATGTTTPARXArgLeuSerAspAsnSerGlnTrp 369
| | | | |
Db 355 CCGCCAGGTGTGAACTGCTGTTTGAACCTGCGCTGAAGCATTAACAGCCAGTGG 296
QY 370 IleglnValSerleuValPheGlnThrleuGLnglnlMerArXAspLyThrProLeuSer 389
| | | | |
Db 295 ATTCAAGTTTCCTGCTCTTCCAGACTTTTACAGCAATGCGATGATTAACAGCGCTGTCA 236
QY 390 LeuAsnThrProProGLYGLuleuValPheGLUATGTTTPARXArgLeuSerAspAsnSerGlnTrp 409
| | | | |
Db 225 TTTAAATACGCCCGCCGAGAGGTGAACCTGACCTGCGAGATGTGAAGAGCAATGCG 176
QY 410 GlnGLYMetCysSerleuAlaGLYPheThrGlnlIleValAsnGLUAlaArglIleProAla 429
| | | | |
Db 175 CAGGCGATGTGTTGTTGCGAGGTTTACGCAATGCTGATGATGAGCAGCATACCGGCG 116
QY 430 CysSerleu 432
| | | | |
Db 115 TGCAGTTTG 107

RESULT 6
CN762997/c 895 bp mRNA linear EST 20-MAY-2004
LOCUS ID0AA12DE01RM1 ApMS Acyrthosiphon pisum cDNA clone ID0AA12DE01 5',
DEFINITION mRNA sequence.
ACCESSION CN762997
VERSION CN762997.1 GI:47536920
KEYWORDS EST.
SOURCE Acyrthosiphon pisum (pea aphid)
ORGANISM Acyrthosiphon pisum (pea aphid)
Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.
1 (bases 1 to 895)
Hunter, W., Martinez-Torres, D., Rahbe, Y., Sabater-Munoz, B.,
Stern, D., Tagu, D. and Wincker, P.
An expressed sequence tags database for the pea aphid Acyrthosiphon
pisum
Unpublished (2004)
Contact: D. Tagu
INRA Rennes
UMR BIO3P, BP 35327, F-35653 Le Rheu Cedex France
Tel: +33.2.23.48.51.65
Fax: +33.2.23.48.51.50
Risk of contamination by bacterial sequences from obligatory
(Buchnera) or facultative endosymbionts. These sequences were
obtained in the frame of the International Consortium of Aphid
Genomics in collaboration with Genoscope
PCR Primers
FORWARD: CAGGAAACAGCTATGACC
Plate: 5 row: C column: 9.
Location/Qualifiers
1..895
/organism="Acyrthosiphon pisum"
/mol_type="mRNA"
/cultivar="developmentstage"
/db_xref="taxon:7029"
/clone="ID0AA12DE01"
/issue_type="whole_insect"
/dev_stage="nymphs and adults (parthenogenetic females)"
/lab_host="XL1-Blue"
/clone_lib="ApMS"
/note="Vector: pBS-SK minus; Site 1: EcoRI; Site 2: XhoI;
Sample name: ID0AA1 ; Plant growth place: Department of

Ecology & Evolutionary Biology, Princeton University ;
Soil conditions: Soil ; Sowing date: 01/06/1999 ;
Harvesting date: 01/06/1999 ; Stress date: no stress ;
Description: Aphids inoculated on one-week old *Vicia faba*
under non-sterile conditions. All parthenogenetic stages
and both winged and wingless adults were collected for
library construction. ; experimental condition: long
photoperiod (16-hr light/8-hr dark at 18 C)"

ORIGIN

Alignment Scores:
Pred. No.: 1,136-74 Length: 895
Score: 838.00 Matches: 159
Percent Similarity: 98.88 Conservative: 0
Best Local Similarity: 98.88 Mismatches: 2
Query Match: 36.48 Indels: 0
DB: 8 Gaps: 0

US-10-601-319-2 (1-440) x CN762997 (1-895)

QY 272 HisAenAlaGlnPheTYrleuLeuGLnArgThrProGLUValAlaArgSerArgAlaThr 291
| | | | |
Db 894 CATTAACGGCAATTTATTTGCTTACACGACGCAAGGTTGCCGCGCGCCGCAAC 835
QY 232 ProLeuAspLeuIleMetAlaAlaLeuThrProHisProProGLnIleGlnAlaTYr 311
| | | | |
Db 834 CCGTATTAGATTATATCAAGACAGCGTTCAGCCCATCCACCCCAAAACAGCGGTAT 775
QY 312 GlyValThrleuProThrsSerValleuPhelelaaglyYHAspThrAsnleuAlaAsn 331
| | | | |
Db 774 GGTGTGACATTACCACTTACGCTGCTTTATGCGGACAGATTAATCTGACCAAT 715
QY 332 leuGLYGLYAlaleuGLuleuAsnTrpThrleuProGLYInProAspAsnThrProPro 351
| | | | |
Db 714 CTCGCGCGCGCACTGAGCTCAACTGACGCTTCCCGGTACGCGGATTAACGCGCGCA 655
QY 352 GlyGLYGLuleuValPheGLUATGTTTPARXArgLeuSerAspAsnSerGlnTrpIlegln 371
| | | | |
Db 654 GGTGTGAACTGCTGTTTGAACGCTGCGCTGCTGCAATGATTAACGCGATTCAG 595
QY 372 ValSerleuValPheGlnThrleuGLnglnlMerArXAspLyThrProLeuSerleuAsn 391
| | | | |
Db 594 GTTTCGCTGCTCTTCCACACTTTACAGAGATGCTGATTAACCCCGCTGCTATTAAT 535
QY 392 ThrProProGLYGLuleuValleuAsnTrpThrleuAlaGLYGLYGLUArgAsnAlaGlnGLY 411
| | | | |
Db 534 ACGCGCGCGGAGAGGTGAACCTGACGAGATGTAAGAGCAAAATGCCAGGAGC 475
QY 412 MetCysSerleuAlaGLYPheThrGlnlIleValAsnGLUAlaArglIleProAlaCysSer 431
| | | | |
Db 474 ATGTGTTGTTGGCAGGTTTACCGCAATCGTGAATGAAGCAGCATACCGCGTGAGT 415
QY 432 Leu 432
| | | | |
Db 414 TTG 412

RESULT 7
CN754382/c 868 bp mRNA linear EST 19-MAY-2004
LOCUS ID0AA12DE01RM1 ApMS Acyrthosiphon pisum cDNA clone ID0AA12DE01 5',
DEFINITION mRNA sequence.
ACCESSION CN754382
VERSION CN754382.1 GI:47519379
KEYWORDS EST.
SOURCE Acyrthosiphon pisum (pea aphid)
ORGANISM Acyrthosiphon pisum (pea aphid)
Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.
1 (bases 1 to 868)
Hunter, W., Martinez-Torres, D., Rahbe, Y., Sabater-Munoz, B.,
Stern, D., Tagu, D. and Wincker, P.
An expressed sequence tags database for the pea aphid Acyrthosiphon
pisum
Unpublished (2004)
Contact: D. Tagu
INRA Rennes
UMR BIO3P, BP 35327, F-35653 Le Rheu Cedex France
Tel: +33.2.23.48.51.65
Fax: +33.2.23.48.51.50
Risk of contamination by bacterial sequences from obligatory
(Buchnera) or facultative endosymbionts. These sequences were
obtained in the frame of the International Consortium of Aphid
Genomics in collaboration with Genoscope
PCR Primers
FORWARD: CAGGAAACAGCTATGACC
Plate: 5 row: C column: 9.
Location/Qualifiers
1..895
/organism="Acyrthosiphon pisum"
/mol_type="mRNA"
/cultivar="developmentstage"
/db_xref="taxon:7029"
/clone="ID0AA12DE01"
/issue_type="whole_insect"
/dev_stage="nymphs and adults (parthenogenetic females)"
/lab_host="XL1-Blue"
/clone_lib="ApMS"
/note="Vector: pBS-SK minus; Site 1: EcoRI; Site 2: XhoI;
Sample name: ID0AA1 ; Plant growth place: Department of

JOURNAL
COMMENT
p1sum
Unpublished (2004)
Contact: D. Tagu
INRA Rennes
UMR BIO3P, BP 35327, F-35653 Le Rheu Cedex France
Tel: +33.2.23.48.51.65
Fax: +33.2.23.48.51.50
Risk of contamination by bacterial sequences from obligatory
(buckera) or facultative endosymbionts. These sequences were
obtained in the frame of the International Consortium of Aphid
Genomics in collaboration with Genoscope

FEATURES
source
Plate: 12 row: E column: 1.
location/Qualifiers

1. 868
/organism="Acyrthosiphon pisum"
/mol_type="mRNA"
/culturvar="developmentstage"
/db_xref="taxon:7029"
/clone="ID0AAA12DE01"
/tissue_type="whole insect"
/dev_stage="nymphs and adults (parthenogenetic females)"
/lab_host="X11-Blue"
/clone_lib="APMS"
/note="Vector: PBS-SK minus; Site 1: EcoRI; Site 2: XhoI;
Sample name: ID0AAA ; Plant growth place: Department of
Ecology & Evolutionary Biology, Princeton University ;
Soil conditions: Soil ; Sowing date: 01/06/1999 ;
Harvesting date: 01/06/1999 ; Stress date: no stress ;
Description: Aphids inoculated on one-week old Vicia faba
under non-sterile conditions. All parthenogenetic stages
and both winged and wingless adults were collected for
library construction. ; experimental condition: long
photoperiod (16-hr light/8-hr dark at 18 c)"

ORIGIN

Alignment Scores:

Pred. No.: 1.99e-67 Length: 868
Score: 767.00 Matches: 149
Percent Similarity: 98.0% Conservative: 0
Best Local Similarity: 98.0% Mismatches: 3
Query Match: 33.3% Indels: 1
DB: 8 Gaps: 0

US-10-601-319-2 (1-440) x CN754382 (1-868)

QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaLeu 300
Db 866 CGCAGCCGAGATT-GCCCGCAGCGCGCCACCCCGTTATTAGATTGATCAAGACAGCG 808
QY 301 LeuThrProHisProProGluIleGlnAlaIleArgValThrLeuProThrSerValLeu 320
Db 807 TTGACGCCCATCCACCGCAAAACAGCGGATGATGATGATTAACCACTTCAGTGTG 748
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeuGlnLeuAsnTrp 340
Db 747 TTTATCGCCGACACGATTAATCTGCAATCTCGCGCGGCACTGAGACTCAACTG 688
QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyLeuValPheGlnArgTrp 360
Db 687 ACGCTTCCCGGTCAACCGGATTAACACGCCGCAAGTGTGAACGTGTTGAAACGCTGG 628
QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 627 CGTGGCGTAAACGATTAACAGCAGGATTCAGGTTTGCATGCTTCCACACTTTACAG 568
QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValValLeuThr 400
Db 567 CAGATGCGTGAATAACCGCGCTGCTAATAATACGCCGCCGAGAGGTAAACTGACC 508
QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
1

Db 507 CTGGCAGGATGTGTAAGACGAAATGCGCAGGCGCATGTGTTCTGTGCGAGGTTTACGCAA 448
QY 421 ILeValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 447 ATGTGAATGAAGACGCAATACCGCGCTGCAATTG 412

RESULT 8
CN757565/c
LOCUS
DEFINITION
ID0AAA1DF09M1 APMS Acyrthosiphon pisum cDNA clone ID0AAA1DF09 5',
mRNA sequence.

ACCESSION
CN757565
VERSION
CN757565.1 GI:47531488
KEYWORDS
EST.
SOURCE
Acyrthosiphon pisum (pea aphid)

ORGANISM
Acyrthosiphon pisum
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
Aphidoidea; Aphididae; Macrosciphini; Acyrthosiphon.

REFERENCE
AUTHORS
Hunter, W., Martinez-Torres, D., Rabhe, Y., Sabater-Munoz, B.,
Stern, D., Tagu, D. and Wincker, P.

TITLE
An expressed sequence tags database for the pea aphid Acyrthosiphon
pisum

JOURNAL
COMMENT
Unpublished (2004)
Contact: D. Tagu
INRA Rennes

UMR BIO3P, BP 35327, F-35653 Le Rheu Cedex France
Tel: +33.2.23.48.51.65
Fax: +33.2.23.48.51.50
Risk of contamination by bacterial sequences from obligatory
(buckera) or facultative endosymbionts. These sequences were
obtained in the frame of the International Consortium of Aphid
Genomics in collaboration with Genoscope

PCR Primers
FORWARD: CAGGAAACAGCTATGACC
Plate: 1 row: F column: 9.
location/Qualifiers

FEATURES

1. 864
/organism="Acyrthosiphon pisum"
/mol_type="mRNA"
/culturvar="developmentstage"
/db_xref="taxon:7029"
/clone="ID0AAA1DF09"
/tissue_type="whole insect"
/dev_stage="nymphs and adults (parthenogenetic females)"
/lab_host="X11-Blue"
/clone_lib="APMS"
/note="Vector: PBS-SK minus; Site 1: EcoRI; Site 2: XhoI;
Sample name: ID0AAA ; Plant growth place: Department of
Ecology & Evolutionary Biology, Princeton University ;
Soil conditions: Soil ; Sowing date: 01/06/1999 ;
Harvesting date: 01/06/1999 ; Stress date: no stress ;
Description: Aphids inoculated on one-week old Vicia faba
under non-sterile conditions. All parthenogenetic stages
and both winged and wingless adults were collected for
library construction. ; experimental condition: long
photoperiod (16-hr light/8-hr dark at 18 c)"

ORIGIN

Alignment Scores:

Pred. No.: 2.09e-66 Length: 864
Score: 757.00 Matches: 146
Percent Similarity: 98.0% Conservative: 0
Best Local Similarity: 98.0% Mismatches: 3
Query Match: 32.9% Indels: 1
DB: 8 Gaps: 0

US-10-601-319-2 (1-440) x CN757565 (1-864)

QY 282 ThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaLeu 301
Db 863 ACGCCAGAGTTGCCCGCAGCGCCACCCCGTTATTAGATTG-ATCAAGACAGCGTTG 805

QY 302 ThrProHisProProGlnInuGlnAlaTyrGlyValThrLeuProThrservalLeuPhe 321
 DB 804 AGCCCATCCACCGGAAAAACAGGGGTATGGTGTACATTACCCACTTCAGTGGCTTT 745
 QY 332 TLeaAglyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTrpThr 341
 DB 744 ATCCCGGACACGATACTATCTGGCAATCTCCGGGGGCGACTGGAGTCAACTGGACG 685
 QY 342 LeuProGlyGlnProAspAsnThrProProGlyGlyGlnLeuValPheGlnuArgTrpArg 361
 DB 684 CTTCCTCCGTCACCGCATACACAGCCGCGAGGTGGTGAACGTGTTTGAACGTGGCGT 625
 QY 362 ArgLeuSerAspAsnSerGlnTrpTyrIleGlnValSerLeuValPheGlnThrLeuGln 381
 DB 624 CGGCTTAAGGATACAGCCAGTGAATCAGGTTTGGCTGCTTCACAGCTTACAGCAG 565
 QY 382 MetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThrLeu 401
 DB 564 ATGGGTGATTAACAGCCGCTGTCTTAATACGCCGCCGAGAGGTGAACCTGACCTTG 505
 QY 402 AlaGlyCysGlnGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGlnIle 421
 DB 504 GCAGATGTGAAGACGAAATGCGAGGGCATGTGTGCTGGCAGGTTTACGCAAAATC 445
 QY 422 ValAsnGluAlaArgIleProAlaCysSerLeu 432
 DB 444 GTGAATGAAGCAGCATACCGGCGTGCAGTTTG 412
 RESULT 9 383 bp mRNA linear EST 18-MAY-2001
 AM036132
 LOCUS
 DEFINITION
 EST274508 tomato seed, TAMU lycopersicon esculentum cDNA clone
 CLEB1E23 similar to periplasmic phosphoanhydride phosphohydrolase
 precursor, putative, mRNA sequence.
 AM036132
 AM036132.1 GI:5894811
 EST.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Lycopersicon esculentum (Solanum lycopersicum)
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 383)
 REFERENCE
 AUTHORS
 Alcalá, J., Vrebalov, J., White, R., Matern, A.L., Vision, T.,
 Holt, I.E., Liang, F., Upton, J., Romning, C.M., Craven, M.B.,
 Fujii, C.Y., Bowman, C.L., Nierman, W., Fraser, C.M., Venter, J.C.,
 Martin, G.B., Tanksley, S.D. and Giovannoni, J.
 Generation of ESTs from tomato seed tissue
 Unpublished (1999)
 TITLE
 JOURNAL
 COMMENT
 Contact: CUGI
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 3 prime sequence.
 FEATURES
 SOURCE
 1. .383
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="CLEB1E23"
 /tissue_type="seeds"
 /dev_stage="quiescent seed"
 /lab_host="XLI-Blue MRP"
 /clone_lib="tomato seed, TAMU"
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
 XhoI; cDNA - Tomato Seed EST library. Directionally cloned
 cDNAs inserted into pBluescript SK(-) at 5' end with
 EcoRI and 3' end with XhoI site."

Alignment Scores:
 Pred. No.: 2,836-55 Length: 383
 Score: 643.00 Matches: 127
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 27.9% Indels: 0
 DB: 7 Gaps: 0
 US-10-601-319-2 (1-440) x AM036132 (1-383)
 QY 117 LysThrGlyGluAlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThr 136
 DB 2 AAAACAGCGAAGCTTGGCCGCGGCTGGCACTGACTGTGCATTAACGTACATACC 61
 QY 137 GlnAlaAspTrpSerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGln 156
 DB 62 CAGGAGATACGTCAGTCCGATCCGATCCGTTATTAATCTCTTAATAAACGGGGTTGGCA 121
 QY 157 LeuAspAsnAlaAsnValThrAspAlaIleLeuSerAlaGlyGlySerIleAlaAsp 176
 DB 122 CTGATTAACGGGAACGTACTGACCGCATCTCTCAGCAGGGGACAGAGGCTCAATTGCTAC 181
 QY 177 PheThrGlyHisArgGlnThrAlaPheArgGlnLeuGluArgValLeuAsnPheProGln 196
 DB 182 TTTCACGGGCAATCGGCAACCGCGTTTGCCTGCAACTGGAACGGGTCTTAATTTTCCGCA 241
 QY 197 SerAsnLeuCysLeuLysArgGlnLysGlnAspGlnSerCysSerLeuThrGlnAlaLeu 216
 DB 242 TCAAACTTGCTGCTTAACGTGAGAAACAGACAGCAAGCTGTTCAATTAAACGAGGCACTTA 301
 QY 217 ProSerGlnLeuLysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAla 236
 DB 302 CCATCGGAACCTCAAGGTGAGCGCGACATGTCTCATTAACCGGTGCGTTAAGCTCGCA 361
 QY 237 SerMetLeuThrGluIlePhe 243
 DB 362 TCATGCTGACCGAGATATTTT 382
 RESULT 10
 AM036134/c
 LOCUS
 DEFINITION
 EST274510 tomato seed, TAMU lycopersicon esculentum cDNA clone
 CLEB1E23 similar to periplasmic phosphoanhydride phosphohydrolase
 precursor, putative, mRNA sequence.
 AM036134
 AM036134.1 GI:5894813
 EST.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Lycopersicon esculentum (Solanum lycopersicum)
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 383)
 REFERENCE
 AUTHORS
 Alcalá, J., Vrebalov, J., White, R., Matern, A.L., Vision, T.,
 Holt, I.E., Liang, F., Upton, J., Romning, C.M., Craven, M.B.,
 Fujii, C.Y., Bowman, C.L., Nierman, W., Fraser, C.M., Venter, J.C.,
 Martin, G.B., Tanksley, S.D. and Giovannoni, J.
 Generation of ESTs from tomato seed tissue
 Unpublished (1999)
 TITLE
 JOURNAL
 COMMENT
 Contact: CUGI
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.
 FEATURES
 SOURCE
 1. .383
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="CLEB1E23"
 /tissue_type="seeds"

```

/dev stage="quiescent seed"
/lab host="XLI-Blue MRP"
/clone lib="tomato seed, TAMU"
/notes="vector: pBluescript SK(-), Site_1: EcoRI, Site_2:
XhoI; CLEB - Tomato Seed EST library. Directionally cloned
cDNAs inserted into pBluescript SK(-) at 5' end with
EcoRI and 3' end with XhoI site."

```

ORIGIN

Alignment Scores:

Pred. No.:	2,83e-55	Length:	383
Score:	643.00	Matches:	127
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	27.9%	Indels:	0
DB:	7	Gaps:	0

US-10-601-319-2 (1-440) x AM036134 (1-383)

Qy 117 LyThrGlyGluAlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThr 136

Db 382 AAAACAGCGCAAGCCTTCGCCGCCGCTGCACTGACATACCGTACATACC 323

Qy 137 GluAlaAspThrSerSerProAspProLeuPheAsnProLeuIleValGlyGln 156

Db 322 CAGGCAATACGTCAGATCCGATCCGTTATTTATCTTAAACCTGCGGTTGCCAA 263

Qy 157 LeuAspSerAlaAsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAsp 176

Db 262 CTGGATTAACGGAAACGTAAGTCAAGCTCTGACAGGAGGAGGCTCAATGCTGAC 203

Qy 177 PheThrGlyHisArgGlnThrAlaPheArgIleuGluArgValLeuAsnProGln 196

Db 202 TTTACCGGCAATCGGCAACCGCTTTCGCAACTGCAACCGGCTTAAATTTCCGCA 143

Qy 197 SerAsnLeuCysLeuValArgGluValGlnAspGlnSerCysSerLeuThrGlnAlaLeu 216

Db 142 TCAAACTTGTCCTTAAACGTGAAGAAACAGGACGAAAGCTGTCATTACCGAGCATTA 83

Qy 217 ProSerGluLeuValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAla 236

Db 82 CCATCGGAACCTCAAGGAGGAGCGGCAACATGTCATTAAACCGGTCGTAAGCTCGCA 23

Qy 237 SerMetLeuThrGluIlePhe 243

Db 22 TCATCTGACGAGATTTT 2

RESULT 11 BE520240/c 393 bp mRNA linear EST 19-MAR-2001

LOCUS M1B12STM Arabidopsis developing seed Arabidopsis thaliana cDNA

DEFINITION clone M1B12 5', mRNA sequence.

ACCESSION BE520240 BE520240.1 GI:9778242

VERSION BE520240.1 GI:9778242

KEYWORDS Arabidopsis thaliana (chale crese)

SOURCE Arabidopsis thaliana

ORGANISM Arabidopsis thaliana (chale crese)

REFERENCE White, J.A., Todd, J., Newman, T., Focke, N., Girke, T., Martinez de

Jarduya, O., Javoricki, J.G., Ohlrogge, J., and Benning, C.

TITLE A new set of Arabidopsis expressed sequence tags from developing

JOURNAL Plant Physiol. 124 (4), 1582-1594 (2000)

COMMENT Contact: Benning, C

Dept. of Biochemistry & Molecular Biology

Michigan State University

224 Biochemistry, Michigan State University, East Lansing, MI

48824, USA

Tel: 517 355 1609

```

Fax: 517 353 9334
Email: benning@msu.edu
Michigan State University DNA Sequencing Facility Arabidopsis
Biological Resource Center, The Ohio State University, 309 Botany &
Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210 USA, FAX:
6142920603 TEL: 6142929371.
location/Qualifiers
1. .393
/organism="Arabidopsis thaliana"
/mol type="mRNA"
/ecotype="Columbia"
/DB_xref="taxon:3702"
/clone="M1B12"
/risue type="seed"
/dev stage="5-13 days after flowering"
/lab host="E.coli"
/clone lib="Arabidopsis developing seed"
/notes="Organ: Developing seed; Vector: pBluescript SK-;
Site_1: EcoRI; Site_2: XhoI"

```

FEATURES

source

ORIGIN

Alignment Scores:

Pred. No.:	4,71e-55	Length:	393
Score:	641.00	Matches:	127
Percent Similarity:	97.7%	Conservative:	1
Best Local Similarity:	96.9%	Mismatches:	3
Query Match:	27.8%	Indels:	0
DB:	7	Gaps:	0

US-10-601-319-2 (1-440) x BE520240 (1-393)

Qy 107 IleIleAlaAspValAspGluArgThrArgLyThrGlyGluAlaPheAlaGlyLeu 126

Db 393 ATTATTCCTATGTCGACGAGCGGACCGTAANAACAGCGGCTTCCGCCGGGANTG 334

Qy 127 AlaProAspCysAlaIleThrValHisThrGlnAlaAspThrSerSerProAspProLeu 146

Db 333 GCACCTGACGTGCATTAACCGTAACATACCCAGGCAATCGTCAGTCCGATCCGTTA 274

Qy 147 PheAspProLeuIleThrGlyValGlyGlnLeuAspAsnAlaAsnValThrAspAlaIle 166

Db 273 TTTATCTCTTAANAACGTGGCGTTTGCCATCGATACCGCAACGTGACGCGCATC 214

Qy 213 CTGACGAGGCGAGGAGGCTCAATGCTGACTTAAACGGGATCGGCAACGGGCTTCCG 154

Qy 167 LeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHisArgGlnThrAlaPheArg 186

Db 213 CTGACGAGGCGAGGAGGCTCAATGCTGACTTAAACGGGATCGGCAACGGGCTTCCG 154

Qy 187 GluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCysLeuValArgGluValGln 206

Db 153 GAACTGGAACGGGCTTAAATTTCCCAATCAAACTTGCCCTTAAACGTGAAGAAACAG 94

Qy 207 AspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeuValSerAlaAspAsn 226

Db 93 GACGAAAGCTGTCATTAAACGACGACCTTCAATCGAATCTCAAGTGAAGCGCGCAAT 34

Qy 227 ValSerLeuThrGlyAlaValSerLeuAlaSer 237

Db 33 GTCTCATTAAGCGCTGCGGTAAAGCTCGCATCA 1

RESULT 12 BU713769 531 bp mRNA linear EST 23-OCT-2003

LOCUS SNAABUG01 Adult SJC 7/94 Schistosoma japonicum cDNA similar to

DEFINITION sp|P07102|PRA_ECOLI PERIPLASMIC APPA PROTEIN PRECURSOR (INCLUDES:

PHOSPHOANHIDRIDE PHOSPHOHIDROLASE (PH 2.5 ACID PHOSPHATASE) (AP);

6-PHATASE), mRNA sequence.

ACCESSION BU713769 BU713769.1 GI:28321125

VERSION BU713769

KEYWORDS Schistosoma japonicum

SOURCE Schistosoma japonicum

ORGANISM Schistosoma japonicum

Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;

Strigeida; Schistosomatidae; Schistosomatidae; Schistosoma.

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
COMMENT
CONTACT
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn.
Location/Qualifiers
1..531
/organism="Schistosoma japonicum"
/mol_type="mRNA"
/strain="Chinese (Anhui) strain"
/db_xref="taxon:6182"
/sex="Male and female"
/tissue_type="Whole body"
/dev_stage="Adult worms"
/lab_host="Mouse and rabbit"
/clone_lib="Adult 5JC 7/94"
/note="Vector: Lambda ZAP-II XR.; Site 1: Ecor I; Site 2:
XhoI I; Several hundred adult Schistosoma japonicum
(Anhui, P.R. China, strain), of mixed sex, were perfused
from the mesenteries of experimentally infected mice and
rabbits at the Queensland Institute of Medical Research,
Brisbane, Australia (QIMR), and stored for several months
in liquid nitrogen. Subsequently, mRNA was isolated at the
QIMR from lysates of these worms by oligo dt
chromatography, using a kit from Pharmacia. The mRNA was
then shipped to Clontech, Palo Alto, CA, USA, who
constructed a cDNA library. First strand synthesis was
primed with an oligo-dt-XhoI-primer and synthesized using
M-MuLV reverse transcriptase. Second strand synthesis was
accomplished with RNase H and T4 DNA polymerase. The
double stranded cDNA was ligated to EcoRI linkers,
digested with EcoRI and XhoI, and ligated into the
phagemid vector lambda ZAP II XR. After construction of
this directional library by Clontech, it was returned to
the QIMR. During analysis of the library at the QIMR, we
have found that a small percentage, 2% to 3%, of the
clones contain inserts that appear to be highly homologous
to sequences from salmonid fishes, as determined by
homology comparisons using BLAST and by Southern
hybridization analysis to genomic DNA from salmon (Sigma
Chemical Co., St. Louis, MO) under stringent washing
conditions. The remainder of the clones appear to contain
S. japonicum sequences."

ORIGIN
Alignment Scores:
Pred. No.: 8.65e-50 Length: 531
Score: 591.50 Matches: 133
Percent Similarity: 79.54 Conservative: 7
Best Local Similarity: 75.64 Mismatches: 30
Query Match: 25.74 Indels: 7
DB: 3 Gaps: 1

US-10-601-319-2 (1-440) x BU713769 (1-531)

QY 224 AlaAspAsnValSerLeuThrglYAlaValSerLeuAlaSerMetLeuThrglUlePhe 243
Db 1 GCCGACAAATGCTCTATTAAACGGGTCGGTAAAGCCCTCGCATCAATCTACGAGATGGGG 60

QY 244 LeuLeuGlnGlnAlaGlnGlyMeCProGluProGlyTTPGlyArgIleThrAspSerHis 263
Db 61 CTCCTGCAACAAGCAACGGAATGCCGAGCCGGGTGGGAGAGATCAACGATTCAAC 120

QY 264 GlnTPAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyLeuLeuGlnArgThrPro 283
Db 121 CAGTGAACACCTTCTAAGTTTGATGCAACGCGCAATTTATTTCTACACGACGCCA 180

QY 284 GluValAlaArgSerArgAlaThrProLeuAspLeuIleMetAlaIleThrPro 303
Db 181 GAGGTGGCCCGGAGCCGCCACCCCGGATTAGATTGATCAAGACAGCGCTTGACGCC 240

QY 304 HisProGlnLysGlnAlaTygIValThrLeuProThrSerValLeu-PheIleAl 323
Db 241 CGTTATATCATTAATGATTTGTCAGTGAAGACAAATTCCTCTGATGACGAAAAAAGG 300

QY 323 AglyHisAspThrAsnLeuAlaAsnLeuGlyGlyValaLeuGluLeuAsnThrThrLeuPr 343
Db 301 CGGCCAGACGCTGAGCTGATTAACCTCGCGCCGACCTGAGCTCAACTGACCGCTGTC 360

QY 343 GGLYGLNProAspAsnThrProProGlyGlyGluLeu-ValPheGluArgArg 362
Db 361 CAGGACGCCCATTAACAGCCCGCAGAGCGGAACCTGATGTTGAACGCTGGCGTCGG 420

QY 363 LeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGlnGlnMet 382
Db 421 CTAACGATTAACAGCACTGATTCAGGTTTCGCTGCTTCCACACTTA-CAGCAGATG 479

QY 383 ArgAspLysThrProLeuSerLeuAsnThrProProGlyGlu 396
Db 480 CGTCGGGGGGTTCCTTGGGG-----CCTTATGGGGAT 512

RESULT 13
LOCUS CN759004/c
DEFINITION ID0AAA24BC04RM1 ApMS Acyrthosiphon pisum cDNA clone ID0AAA24BC04
5', mRNA sequence.
ACCESSION CN759004
VERSION CN759004.1 GI:47532927
KEYWORDS EST.
SOURCE Acyrthosiphon pisum (pea aphid)
ORGANISM Acyrthosiphon pisum
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
Aphididae; Aphidinae; Macrosiphini; Acyrthosiphon.
1 (bases 1 to 706)
Hunter, W., Martinez-Torres, D., Rahbe, Y., Sabater-Munoz, B.,
Stern, D., Tagu, D. and Winkler, P.
An expressed sequence tags database for the pea aphid Acyrthosiphon
pisum
Unpublished (2004)
JOURNAL
COMMENT
INRA Rennes
UMR BIO3P, BP 35327, F-35653 Le Rheu Cedex France
Tel: +33.2.23.48.51.65
Fax: +33.2.23.48.51.50
Risk of contamination by bacterial sequences from obligatory
(Buchnera) or facultative endosymbionts. These sequences were
obtained in the frame of the International Consortium of Aphid
Genomics in collaboration with Genoscope
PCR Primers
FORWARD: CAGGAACAGCTATGACC
Plate: 24 row: C column: 4.
FEATURES
source
1..706
Location/Qualifiers
/organism="Acyrthosiphon pisum"
/mol_type="mRNA"
/culivar="developmentstage"
/db_xref="taxon:7029"
/clone="ID0AAA24BC04"
/tissue_type="Whole Insect"
/dev_stage="nymphs and adults (parthenogenetic females)"
/lab_host="XII-Blue"
/clone_lib="ApMS"
/note="Vector: pBS-SK minus, Site 1: EcoRI, Site 2: XhoI;
Sample name: ID0AAA ; Plant growth place: Department of

Ecology & Evolutionary Biology, Princeton University ;
Soil conditions: Soil ; Sowing date: 01/06/1999 ;
Harvesting date: 01/06/1999 ; Stress date: no stress ;
Description: Aphids inoculated on one-week old *Vicia faba*
under non-sterile conditions. All parthenogenetic stages
and both winged and wingless adults were collected for
library construction. ; experimental condition: long
photoperiod (16-hr light/8-hr dark at 18 °C)

ORIGIN

Alignment Scores:

Pred. No.:	1.73e-42	Length:	706
Score:	522.00	Matches:	98
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	22.7%	Indels:	0
DB:	8	Gaps:	0

US-10-601-319-2 (1-440) x CN759004 (1-706)

Qy 335 AAlaenGluLeuAaNTpThleuProglYglInProaSpaenThProProglYglYglu 354
Db 705 GCACTGGAGCTCACTGACCGCTTCCGGTCAACCGGATTAACGCCCGCAGTGTGAA 646
Qy 355 LeuValPhegluAgtTpaArgleuSerApsaenSerGlntPileglInValSerleu 374
Db 645 CTGGTGTGTAACCGTGGCGCTCGGCTAACGATTAACGACGTGGATTCAAGTTTCGCTG 586
Qy 375 ValPheglntThleuGlntGlnMetCArgApylThrProleuSerleuAaenThrProPro 394
Db 585 GTCCTCCAGACTTACACAGCAGATGGTGATAAAGCGCGCTGTCTTAATAACGCCGCC 526
Qy 395 GLYglVallylsleuThleuAlaglyCysegluGluaArgaAnaIaglnGlyMetCyseSer 414
Db 525 GGAAGAGTGAACGACCCCTGGCAGAGATGTGAAGAAGCCAAATGCCAGAGGATGTTCG 466
Qy 415 LeuAlaglyPheThrglnIleValAsnGluAlaArgIleProAlaCyseSerleu 432
Db 465 TTGGAGGTTTATGACCAATCGTGAATGAAGACGCATACCGCGCTGCAGTTTG 412

RESULT 14
AA545747/c 354 bp mRNA linear EST 12-MAY-1999
LOCUS HBMSF1B4-REV Human Bone Marrow Stromal Fibroblast Homo sapiens cDNA
DEFINITION
ACCESSION AA545747
VERSION AA545747.1 GI:2307026
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 354)

REFERENCE
AUTHORS Jia,L., Robey,P., Young,M., Schuler,G., Powell,J., Yang,L.,
Lennon,G., Hillier,L., Allen,M., Bowles,L., Geisel,S., Kucaba,T.,
Marrin,M., Martin,J., Steptoe,M., Tan,F., Teising,B., Bowers,Y.,
Wylie,T., Waterston,R., Wilson,R. and Francomano,C.
TITLE WashU-MGB/NHGRI EST Project
JOURNAL
COMMENT Unpublished (1997)
CONTACT: Libin Jia
National Human Genome Research Institute
10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
Tel: 301-402-4877
Fax: 301-496-7157
Email: libin@helix.nih.gov
Seq primer: M13 Reverse.

FEATURES

1..354
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

/clone="HBMSF1B4"
/sex="Male and Female"
/tissue_type="bone marrow"
/cell_type="stromal fibroblast"
/dev_stage="mixed"
/lib_host="X11-Blue"
/clone_lib="Human Bone Marrow Stromal Fibroblast"
/notes="Vector: pBluescript; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Alignment Scores:

Pred. No.:	7.2e-42	Length:	354
Score:	511.50	Matches:	112
Percent Similarity:	95.0%	Conservative:	1
Best Local Similarity:	94.1%	Mismatches:	5
Query Match:	22.2%	Indels:	4
DB:	1	Gaps:	0

US-10-601-319-2 (1-440) x AA545747 (1-354)

Qy 166 lIleuSerAArgAlaglySerIleAlaSpheThrglyHisArgIlnThr-Alaph 185
Db 352 GTCCTTCAGCAGGCGAGAGGTCA-ATTGCTGA-TTTACGGGCATCGGCAACGGCGTT 295
Qy 185 eArgIuLeuGluuArgValleuAenPheProglInSerAaenCysleuYarGluJ 205
Db 294 TCGGAACTGGAAGCGGCTGTT-AAATTTCGCAATCAACTGTGCTTAAACGTGAGAA 236
Qy 205 eGlnAapGluSerCyseSerleuThrglnAlaIleuProSerGluLeuYValSerAla 225
Db 235 ACAGAGAGAAAGCGTTCATTACGACGACATTAACATCGAACTCAAGGAGACGCCGA 176
Qy 225 PaaValSerleuThrglyAlaValSerleuAlaSerMetleuThrglyIlePheleu 245
Db 175 CAATGTTTCATTAAACCGGTGGTGAACCTCGATCAATCTAGCGAGATATTTCCT 116
Qy 245 uGlnAlaIaglnGlyMetProgluProglYtTpglyArgIleThraSpSerhsglntr 265
Db 115 GCAACAGACACAGGAGATGCCGAGCGGGGTGGGAAGATCACCGATTACACACAGTG 56
Qy 265 PaaThrleuLeuSerleuHisAaenAlaglnPheTyIleuGlnInrghrPro 283
Db 55 GAAACACCTTGCTTAATGTCATTAACGCGCAATTTATTGCTACACAGCACGCCA 1

RESULT 15
B619443 707 bp mRNA linear EST 01-OCT-2003
LOCUS B619443
DEFINITION
Xenopus laevis CDNA clone X1189d24 5', mRNA sequence.
ACCESSION B619443
VERSION B619443.1 GI:37258203
KEYWORDS
SOURCE EST.
ORGANISM Xenopus laevis (African clawed frog)
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Xenopus.
1 (bases 1 to 707)

REFERENCE
AUTHORS Kohara,Y.
TITLE Expressed genes in X. laevis embryo
JOURNAL
COMMENT Unpublished (2001)
CONTACT: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tahin@genes.nig.ac.jp
The information of this clone is available through the following
URL.
http://xenopus.nibb.ac.jp.
location/Qualifiers

FEATURES

1..707
location/Qualifiers
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:9606"

source

```
1..707
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="XL89d24"
/tissue_type="whole embryo"
/dev_stage="stage 10.5"
/clone_lib="NIBB Mochi normalized Xenopus early gastrula
library"
```

ORIGIN

Alignment Scores:

Pred. No.:	2,666-34	Length:	707
Score:	442.00	Matches:	83
Percent Similarity:	97.6%	Conservative:	0
Best Local Similarity:	97.6%	Mismatches:	2
Query Match:	19.2%	Indels:	0
DB:	2	Gaps:	0

US-10-601-319-2 (1-440) x B619443 (1-707)

```
OY      251 MetProGluProGlyTyrGlyValArgIleThrAspSerHisGlnTrpAsnThrLeuLeuSer 270
      2 ATGCCGAGACCGGGGTGGGAGAGATCACCGATTCAACCACTGGAGACACTTGCTAAGT 61
OY      271 LeuHisAsnAlaGlnPheTyrLeuLeuGlnArgThrProGluValAlaArgSerArgAla 290
      62 TTGCAATAACGGCGCAATTATTTGCTAACACGACGCGAGAGTTGCCCGCAGCCGCC 121
OY      291 ThrProLeuLeuAspLeuIleMetAlaAlaLeuThrProHisProProGlnIlysglnAla 310
      122 ACCCGCTATTGATTGATTCAGACAGCGGTTGACGCCCATCCACCGCAAAACAGGGCG 181
OY      311 TyrGlyValThrLeuProThrseryAlaLeupheileAlaGlyHisAspThrAsnLeuAla 330
      182 TATGGTGTGACATTAACCACTTCAGTGTCTTTATCGCCGACACGATACTAATCTGGCA 241
OY      331 AsnLeuGlyGlyAla 335
      242 AATCTCGGGGGCGCA 256
DB
```

Search completed: June 14, 2006, 15:26:30
Job time : 4992.33 secs

THIS PAGE LEFT BLANK

```

1  FILE REFERENCE: MOA-A0205Y1-US
2  CURRENT APPLICATION NUMBER: US/10/449,902
3  CURRENT FILING DATE: 2003-05-29
4  PRIOR APPLICATION NUMBER: JP 2002-203269
5  PRIOR FILING DATE: 2002-05-30
6  PRIOR APPLICATION NUMBER: JP 2002-383870
7  PRIOR FILING DATE: 2002-12-11
8  NUMBER OF SEQ ID NOS: 56791
9  SOFTWARE: PatentIn Ver. 2.1
10 SEQ ID NO 27858
11
12 LENGTH: 3731
13
14 TYPE: DNA
15
16 ORGANISM: Oryza sativa
17
18 PUBLICATION INFORMATION:
19
20 DATABASE ACCESSION NUMBER: AK102280
21
22 DATABASE ENTRY DATE: 2002-08-28
23
24 US-10-449-902-27838

```

Alignment Scores:	
Pred. No.:	18.6
Score:	95.0%
Percent Similarity:	34.6%
Best Local Similarity:	20.9%
Query Match:	4.1%
OB:	6
	Length: 3731
	Matches: 80
	Conservative: 52
	Mismatches: 150
	Indels: 100
	Gaps: 16

US-10-601-319-2 (1-440) X US-10-449-902-27838 (1-3731)

QY	120	GIuValaPhealalaglyLeuAlaProharyCybalalietPrValHlethr-----	136
Db	1999	GATGCTTCCTTACTTGGTGTGTGCAACTGATCATGTATCAAAAGCATATGCTGGACATGCTTTC	2058
QY	136	-----	136
Db	2059	CAGACCCAAACAAAGAGCTCTCCCAAGATATTCAGTTCCAGACGGCAGAAACATA	2118
QY	137	-----GIuAlaapThrSerSerProharyProleuPheanProleuylThylVal	154
Db	2119	GGTGGCAGAGCTATATACCTTCTCCCAAGTCAGATTTTGGCAGCAAGTGCACCTCAGTTT	2178
QY	155	Cysglu-----LeuaspAsnalasanyalThraPalalileuSerArglagly	171
Db	2179	CAGCAGCAGCAATTCCTCCAAACATCAACGAGAC-----	2214
QY	172	GlyserlilealaaPheThrlGlyHleArgGlnThralaPhe-----ArgGluLeuGlu	189
Db	2215	-----ACATTCCAAGGGCAGATCTCAGCTGAGTCTCTTCAGACAGCTCCAA	2262
QY	190	ArgValLeuapPheProGlnSerAsnleuCysleuylbArgGluGlnaPgluSer	209
Db	2263	CGCTGCCAGTGTCTTATCTGACGAGAAGCCACAGCTGCACAAACCCAGACACACAGAAAGAA	2322
QY	210	CysSerleuthrGlnAlaLeuProSerGluLeu-----	220
Db	2323	TCACAGCAACAAACAGCAGCAGCAATCGCAGTGTATGCAAGTCCACAAACATCAGCAATG	2382
QY	221	-----LysValSerAlaaspAsnalSerleuthrGlyAlavalSerleuAlaSer	237
Db	2383	CAACAGCAAGAAACATGACCAACTATCATGTGTACCTTAATGATATGTGACACCGTTTCTT	2442
QY	238	MetLeuthrGlu-----IlePheLeuGlu	246
Db	2443	CAACTCTCTCTCTCTCTCTCACTTCACACTTACATGACATGCAACAAAGATTAACATTCTCC	2502
QY	247	GlnAlaGlnGlyMetProGluProGlyTTrpGlyArgIleThraSerPheIsglnTrrAsn	266
Db	2503	CAGCCACAGAGCTATCCAGACACAGATGATGAGCTCATTTATCACCATCC-----AAC	2553
QY	267	ThrleuSerleuHleAsnalGlnPheTyrlleuGlnaGlnaGlnTrrPro-----	283
Db	2554	ACA---TCCACCATGCTATATGATTCATTTGAGGCCATTTTCATCGGAAGCACCTTGCGACCTTA	2610
QY	284	GluValAlaArgSerArgAlaThrProleuLeuAsp-----Leuile	297

Db 2611 AGCATGCCAAGCCGAGTACAGTACCTGCTCCCTGATACCAGATGATCATGAAAGCGATTGCA 26700
::: ||| ||| |||::: |||
Qy MetalalaleuThrProHisProProGlnIleuGlnIleValIleuValIleuProThr 317
::: ||| ||| |||::: |||
Db 2671 GTGAGAGCTTTGCTTCCTCCTCTCGTCTCTCAGGTTCATCGCAGATGAGACGACTTGACTCA 2730
::: ||| ||| |||::: |||
Qy 318 SerValIleuPheIleValIleHisAspThrValIleuValIleuValIleuValIleuGlu 337
::: ||| ||| |||::: |||
Db 2731 ACACACACCTAGTATACCA---CAAGAGCTGCGTGGCGCGCTTCTCTGGA----- 2778
::: ||| ||| |||::: |||
Qy 338 LeuValIleuThrIleuProGlnIleuProAspAsnThrProProGlnIleuValIleuValPhe 357
::: ||| ||| |||::: |||
Db 2779 ---AGAGAGAGCTTGGTGATGCAAGATGGGAAAGCTGATCTCTCAAAACCAATCTGTATT 2835
::: ||| ||| |||::: |||
Qy 358 GluArgTrpArgIleuSerAspAsnSerGlnIleuValIleuValIleuValIleuValPheGln 377
::: ||| ||| |||::: |||
Db 2836 -----GGCGTTACATGATTCATCAGTCACCTTCAATGCAAGGTGGCATCCCA 2883
::: ||| ||| |||::: |||
Qy 378 ThrIleuGlnIleuValIleuValIleuValIleuValIleuValIleuValIleuValIleuVal 388
::: ||| ||| |||::: |||
Db 2884 AGGCTCCAGGGGTGAGATGATTGACGGGCAATTCCTATTCCACCTCCAAATTTCTGTAGC 2943
::: ||| ||| |||::: |||
Qy 389 ---SerIleuValIleuValIleuValIleuValIleuValIleuValIleuValIleuValIleuVal 407
::: ||| ||| |||::: |||
Db 2944 CCGTGGCAGAAATGATTTCCCTTGGATGCAACATTAAGTGGTGGAGACTGTTAGATGAA 3003
::: ||| ||| |||::: |||
Qy 408 AsnAlaIleuGlnIleuValIleuValIleuValIleuValIleuValIleuValIleuValIleuVal 427
::: ||| ||| |||::: |||
Db 3004 TCAGAGATATGCGCGCTTCA---CAAAATTCGATCAAGATTAAT-----CGACCA 3054
::: ||| ||| |||::: |||
Qy 428 ProAla 429
||| ||| |||
Db 3055 CCTGCG 3060
||| ||| |||

Search completed: June 14, 2006, 14:24:53
Job time : 59.3073 secs

THIS PAGE LEFT BLANK

Db 1864 GCTGAACAAGAACTGGGGATTGAGGTCAGACCCC-----GTGCCCCCGTGAAGTTGCA 1917
Qy 352 -GlyGlyGluLeuValPheGluArgTTPArgArgLeuSerAspAsnSerGlnTrpIleG1 371
Db 1918 GGGCGCAATCTCGGTGGTTTCTCGGAATCACTGCTCCGGTGGGTGACTATGGTGAC 1977
Qy 371 nValSerLeuValPheGlnTrpLeuGlnGlnMetArgAspLysThr 386
Db 1978 AATCATCATGACGAATAATGACCAAAACAATGATGACTATCAACC 2023

RESULT 14

US-10-449-902-24588
; Sequence 24588, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 24588
; LENGTH: 1488
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE accession number: AK100030
; DATABASE entry date: 2002-08-28
US-10-449-902-24588

Alignment Scores:

Pred. No.: 5.04 Length: 1488
Score: 95.00 Matches: 92
Percent Similarity: 33.9% Conservative: 55
Best Local Similarity: 21.2% Mismatches: 133
Query Match: 4.1% Indels: 154
DB: 6 Gaps: 24

US-10-601-319-2 (1-440) x US-10-449-902-24588 (1-1488)

Qy 9 LeuSerLeuLeuIleProLeuThrProGlnSerAlaPheAlaGlnSerGluPro----- 26
Db 30 CTCGATCCCTCTCTCCCTGTCTCCCGATCCCTTCTCTCTCCCTTCTCCGGCGCG 89
Qy 27 -----GluLeuLysLeuGluSerValIleVal 36
Db 90 CTTTCTCGCGCGCTCCCTTCGCGCGCTCCCTCCCTCCCTACATCCCGACCTCC 149
Qy 37 SerArgHisGlyValArgAlaProThrLysAlaThrGlnLeuMetGlnAspVal----- 54
Db 150 ACCTCCACCGCCCTCGCGCTCTCCGGAGGAGGACAGACCCCTACGACCTTCCATT 209
Qy 55 -----ThrProAspAlaTrpProThrTrpProValLysLeuGlyTrpLeuThrPro 71
Db 210 CGTCGCGCGCGCGCTCGCGCGCTCTCCAGCGCGACATCGAGGTAAGAGCTCCCGCCC 269
Qy 72 ArgGlyGlyGlu-----LeuIleAlaTyrLeuGlyHisTyrGlnArgGlnArgLeuVal 89
Db 270 TTCTCCCGCGGACAGCTGTCTTATGATTTCTCCCCCATTTG-----CGTCGCGGATTGGCC 326
Qy 90 AlaAspGlyLeuLeuAla----- 95
Db 327 GGCATGATGATTACGGCTCTCTTCCTCCGGGGCGCGGACACACATCTTCGACGCCAT 386
Qy 96 ---LysLysGlyCys-----ProGlnSerGlyGlnVal----- 105

Db 387 CGCCGTCGCGCGCTCCGACACCCCGCGCGCTCAGTCCAGCGCGGAGCCCATCGCCA 446
Qy 106 AlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGluAlaPheAlaAlaGly 125
Db 447 GCGCTTCTACCGCGCTACCGTGGCGGTGCTCCACCGGGGCGACCCGCCACCGAGCA 506
Qy 126 LeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThrSerSerProAspPro 145
Db 507 CGCCCGCGCGAG-----CTGCTCCACCCGAGGCGCGCTAGCTGCTCCCT----- 551
Qy 146 LeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAlaAsnValThrAspAla 165
Db 552 -----AGCATCTGTAGTCCGACCGCGCCAATGTC----- 581
Qy 166 IleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGly-----HisArgGln 182
Db 582 -----ATCGCGGACGACGCGCGCTCCCTTGCCTATAGGAC 617
Qy 183 ThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCysLeuLys 202
Db 618 CGCGTCGCGCGGACCGAGCGGCATC-----AAGGCCGTCCTCTCTC 659
Qy 203 ArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuPro-----SerGluLeu 220
Db 660 AAGCAACAAGAAAGTCGGAGCGAGTGTCTCGAGCTGCTCGCGCTCTGCAGCAGCTG 719
Qy 221 LysValSerAlaAspAsnValSerLeuThr-----GlyAlaValSer----- 234
Db 720 GAGCTCAGCGTGGACACCTTGACGGTTACTGAGATTGGGAAGCGGTGAGCAGCTACCGG 779
Qy 235 -----LeuAlaSerMetLeuThrGluIlePheLeuLeu 245
Db 780 AAGCAACAATCCAGCAGATTGTCACCTTGTTCATTTGCTCATCGAA----- 827
Qy 246 GlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThrAspSerHisGlnTrp 265
Db 828 -----CGTTGGAAGCGCATCTCGAT-----GAGTGG 854
Qy 266 AsnThr-----LeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGlnArgThr 282
Db 855 ATGACGACGTGAGACGCCCATTTAGTGCATTTGGATGGAATTTGTTGTGCAAGTGC 914
Qy 283 ProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAlaLeuThr 302
Db 915 AGCCAA-----AAGCACCAGCATCTCTCTGCTTA----- 944
Qy 303 ProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeuPheIle 322
Db 945 -----GCACAACAGATTGCTTT----- 962
Qy 323 AlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrpThrLeu 342
Db 963 -----CTCTAAGCATAGTGGCGCTGAA--GATGAGCTTGG----- 997
Qy 343 ProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrpArgArg 362
Db 998 -----CATCCAGATGGTGAATTAATTTTTCATGTACAGT--- 1033
Qy 363 LeuSerAspAsnSerGlnTrpIleGlnValSerLeuVal 375
Db 1034 ---GCTGAATAACGGCGAC---TCTCAAGTTTCACTGATG 1066

RESULT 15

US-10-449-902-27838
; Sequence 27838, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF

Db 1205 GTGATAGGCTCCTCAGAGTCGGATCTATCATCATCTCACCTCTTCCATGCTATCATCA 1264
 QY 377 InThrLeuGlnGlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluV 397
 Db 1265 AGGAGACATTCGCTCATCATCTCAACACCGCTCTCGCTG----- 1305
 QY 397 alLysLeuThrLeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaG 417
 Db 1306 ----- 1339
 QY 417 lyPheThrGlnIleValAsnGluAlaArgIlePro 428
 Db 1340 GCTAC-----CGTATCCCC 1353

RESULT 13

US-10-449-902-26929
 ; Sequence 26929, Application US/10449902
 ; Publication No. US20060123505A1
 ; GENERAL INFORMATION:
 ; APPLICANT: National Institute of Agrobiological Sciences.
 ; APPLICANT: Bio-oriented Technology Research Advancement Institution.
 ; APPLICANT: The Institute of Physical and Chemical Research.
 ; APPLICANT: Foundation for Advancement of International Science.
 ; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
 ; FILE REFERENCE: MOA-A0205V1-US
 ; CURRENT APPLICATION NUMBER: US/10/449,902
 ; CURRENT FILING DATE: 2003-05-29
 ; PRIOR APPLICATION NUMBER: JP 2002-203269
 ; PRIOR FILING DATE: 2002-05-30
 ; PRIOR APPLICATION NUMBER: JP 2002-383870
 ; PRIOR FILING DATE: 2002-12-11
 ; NUMBER OF SEQ ID NOS: 56791
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 26929
 ; LENGTH: 2293
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 ; PUBLICATION INFORMATION:
 ; DATABASE ACCESSION NUMBER: AKI02371
 ; DATABASE ENTRY DATE: 2002-08-28
 US-10-449-902-26929

Alignment Scores:

Pred. No.:	Score:	8.35	Length:	2293
Percent Similarity:	95.50	Matches:	99	
Best Local Similarity:	32.6%	Conservative:	69	
Query Match:	19.2%	Mismatches:	180	
DB:	4.1%	Indels:	168	
	6	Gaps:	26	

US-10-601-319-2 (1-440) x US-10-449-902-26929 (1-2293)

QY 7 ProPheLeuSerLeuLeuIleProLeuThrProGlnSerAlaPheAlaGlnSerGluPro 26
 Db 574 CTGGGACATCATCTATTCTCTCAATTGATCCCAACATCTAAAGCTCAAAATGAGATG 633
 QY 27 GluLeuLysLeuGluSerValIleValSerValIleValSerArgHisGlyValArg---AlaProThr 45
 Db 634 AGGCGCATTTTGGATCAAAAGTAGTGATTCACCTTGAATAATCAACGGAATGTGCGTAGC 693
 QY 46 LysAlaThrGlnLeuMetGlnAsp----- 53
 Db 694 AGTTCACATCAAGGCTACGAGGTGTAGGCGCGTTGCGCCCAATCCCAAGAAAAACCCCTT 753
 QY 54 ---ValThrProAspAla---TriProThrTriProValLysLeuGlyTriProLeuThrPro 71
 Db 754 CTGTGTACTCCATCTAGTACTGGCCACCATGGGATAAATCAATGCTCAATG----- 804
 QY 72 ArgGlyGlyGluLeuIleAlaTyrrLeuGlyHisTyrrGlnArgGlnArgLeuValAlaAsp 91
 Db 805 -----GATATTGTGGAGACAAAGAGTGGTTTCAACTATTTTCAGGTACATATATGAT 855
 QY 92 GlyLeuLeuAlaLysGlyCysProGlnSerGlnValAlaIleAlaAspVal 111

Db 856 -----CCTTCTGCAGTCATGTGCAAGACTTG 882
 QY 112 AspGluArgThrArgLysThrGlyGlu-----AlaPheAlaAlaGlyLeuAla----- 127
 Db 883 TTTGAAGTTGCCAAATCTGCAAAATGATCTCAATGCTATTGCGCTATATTAGCAAGTAC 942
 QY 128 -----ProAspCysAlaIleThrVal-----HisThrGlnAlaAsp 139
 Db 943 CCTTATATCCAGAGATCATTTGTTAACTTTCTGAACTTTTCAAAATATTCTGGGAACAT 1002
 QY 140 ThrSerSerProAsp-----Pro 145
 Db 1003 CAATCGTCAGCAGATGCTGTTGAGAAGTGTCTATTGCTCTGGAGTGTGCTGGCATCCA 1062
 QY 146 LeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAlaAsnValThrAsp--- 164
 Db 1063 TTGTTTAGCCCACTGCACAGCAAC---TGCCAGCTGAAATATAGCCATGACACAAACAAG 1119
 QY 165 -----AlaIleLeuSerArgAlaGly----- 171
 Db 1120 CCATTCTTTACTGCGCTTTTGTAGTCACATGAAAAATTTGGATAGACGTGGCTGCCATCGG 1179
 QY 172 -----GlySer 173
 Db 1180 TCTGTTTAGAGTCTGCAAAATTCCTTCTTTCATTGATTCCTGATGATCCAAAGGTGCT 1239
 QY 174 Ile-----AlaAspPheThrGlyHisArgGlnThrAlaPheArgGluLeuGluArgVal 191
 Db 1240 CTATTCTGCATTGATTCTTGTCTAAGATCACAAACAGTACAAATGGTTGGACAGTTT 1299
 QY 192 LeuAsn-----PheProGlnSerAsnLeuCysLeu 201
 Db 1300 GCAGAAGAGTACCAGTGTGATACTCTTGTGTTATTCCCTTAATTTCTCATTTTCCCTT 1359
 QY 202 -----LysArgGluLysGln 206
 Db 1360 GCCATTGACGGTCTTAGTGGGTGATGCCAAATAATGGGGAGGCTTTCGACCATTTCT 1419
 QY 207 AspGluSerCysSer-----LeuThrGlnAlaLeu-----ProSerGluLeu 220
 Db 1420 GACAAGTCAACATCTGTAGATCTCATGAAGCAAGCATTTGATGCTTCATCTTTTGGTGTG 1479
 QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 Db 1480 CGCAAGATTGTGCAAGGCTCTTTGAAAGACTCC---TCGTGGACCCAAATACTCAGG 1536
 QY 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
 Db 1537 AATGTGTTCTTT-----GGATCAGCAAGCCAGGA----- 1566
 QY 261 AspSerHisGlnTriProAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrrLeuLeuGln 280
 Db 1567 ---AGCCCTTCCCTTGAGCATATGATTAGCATATATGTTGAACGCCATTATCATGTGG 1623
 QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIle-----Met 298
 Db 1624 AGGTTCCAGAACTGCAGAAATTTACTCAAAGATGCTGCTCTTTTGGTGTGTAATCACTA 1683
 QY 299 Ala-AlaLeuThrProHisPro-----ProGlnLysGlnAlaTyrrG1 312
 Db 1684 AAACAGGATACAGAGAGGCCAGACTGGGCATGTGTAGAAAAGAAAGCATTTCTTTCA 1743
 QY 312 yValThrLeuProThrSerValLeuPheIle-----AlaGlyHisAspTh 327
 Db 1744 GATAAGAATGAGTACTCTCATCTTCTTGTGTTTTCAGACTTCTCCGACACAAACACCATCTC 1803
 QY 327 rAsnLeuAlaAsnLeuGlyGlyAla----- 335
 Db 1804 CCCCTGAAGAATTTGCGGCCCATTTCAATGGTGTGCTCTCTGGAATGGCAGATGATGCGCCA 1863
 QY 336 -LeuGluLeuAsnTriProLeuProGlyGlnProAspAsnThrProPro----- 351

Db 729 GGCCTGTCGGAGGCCACAGGTTCCGCTGTCATCGGAGAGCCAGATCGTAGGCATC 788
 Qy 360 TrpArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeu 379
 Db 789 AGGACGACAGACAGAGCTCGAGTCGGAGTGGCTGCAGGTG-----CTT 830
 Qy 380 GlnGlnMetArgAspLys-----ThrProLeuSerLeuAsnThrProProGlyGluVal 397
 Db 831 GGCAGCTCTACAGAGAGCGGTGATCCGATCGGCTGTTCCCTCCGCG-----CCC 884
 Qy 398 LysLeuThrLeuAlaGlyCysGlu-----GluArgAsnAlaGlnGly 411
 Db 885 ACACAAGACATCGCGCGCACAGAGCGACTTTGCGGTGGCTGGACACAGAGGCTCAGGCG 944
 Qy 412 MetCysSerLeuAlaGlyPheThrGlnIleValAsnGluAlaArgIleProAlaCysSer 431
 Db 945 TCCGTGCTGTACGCGGCTTCGGC-----AGCAGGCGAGACTGACGAGCGCGCAG 995
 Qy 432 LeuArgSer 434
 Db 996 CTCACAAACA 1004

RESULT 12

US-10-449-902-20112
 ; Sequence 20112, Application US/10449902
 ; Publication No. US20060123505A1
 ; GENERAL INFORMATION:
 ; APPLICANT: National Institute of Agrobiological Sciences.
 ; APPLICANT: Bio-oriented Technology Research Advancement Institution.
 ; APPLICANT: The Institute of Physical and Chemical Research.
 ; APPLICANT: Foundation for Advancement of International Science.
 ; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
 ; FILE REFERENCE: MOA-A0205Y1-US
 ; CURRENT APPLICATION NUMBER: US/10/449,902
 ; CURRENT FILING DATE: 2003-05-29
 ; PRIOR APPLICATION NUMBER: JP 2002-203269
 ; PRIOR FILING DATE: 2002-05-30
 ; PRIOR APPLICATION NUMBER: JP 2002-383870
 ; PRIOR FILING DATE: 2002-12-11
 ; NUMBER OF SEQ ID NOS: 56791
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 20112
 ; LENGTH: 2010
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 ; PUBLICATION INFORMATION:
 ; DATABASE ACCESSION NUMBER: AK070442
 ; DATABASE ENTRY DATE: 2001-12-06
 US-10-449-902-20112

Alignment Scores:

Pred. No.: 6.21 Length: 2010
 Score: 96.00 Matches: 112
 Percent Similarity: 34.3% Conservative: 57
 Best Local Similarity: 22.7% Mismatches: 174
 Query Match: 4.2% Indels: 152
 DB: 6 Gaps: 24

US-10-601-319-2 (1-440) x US-10-449-902-20112 (1-2010)

Qy 11 LeuLeuIleProLeuThrProGlnSerAlaPheAlaGlnSerGluProGluLeuLysLeu 30
 Db 106 CTAGTCTCTACCA---ACAATGGAGGTGCGCCCATGGAGATCTTACTCATTTGCTCTC 162
 Qy 31 GluSerVal-----ValIleValSerArgHisGly--- 40
 Db 163 ACCACCGTGGCTCTCTCGTCATCGTGTGTACGCCCTGGTCTTCTCCGCGCGCGGAAG 222
 Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 Db 223 GCGCGTGGCGCG-----CTGCGGTGGCGCTGGCCCCCAGGGATGGCGG 267
 Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGly----- 74

Db 268 GTG-----CTGGCAACCTGCGCCAGCTGGCGGGAGAGCCAGACCGACGCTG 315
 Qy 75 ---GluLeuIleAlaTyrLeuGlyHisTyrGlnArgGlnValAlaAspGlyLeu 93
 Db 316 CACGAGATGACCAAGGTGTACGCGCTGATCGGCTCGGTTCCGGAGCTCCGACGCTG 375
 Qy 94 LeuAlaLysGlyCysProGlnSerGlyGlnValAlaIleIleAlaAspValAspGlu 113
 Db 376 GTGGTCGCGGCTCGCGCGCGGTGGCGGCGCATCTCTCCGACCCACCATGCAAC--- 432
 Qy 114 ArgThrArgLysThrGlyGluAlaPhe-AlaAlaGlyLeuAlaProAspCysAlaIleTh 133
 Db 433 -----TTGACGAGCGCGCCACGCAACTCCGCGCGCGGACAC 468
 Qy 133 rValHisThrGlnAlaAspThrSerProAspProLeuPheAsnProLeuLysThrG1 153
 Db 469 ATGGCGTACAAACGCGCGGACGTCGTGTTG-GGGCGG---TACGGGCGCGGTGGCGCG 524
 Qy 153 yValCysGlnLeuAspAsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlyGlySe 173
 Db 525 CATGCGGAAGATTGCGCGCTCAACCTCTCTCCGCG-----CGCGCGCTCGACGA 575
 Qy 173 rIleAlaAspPheThrGlyHisArgGlnThrAlaPheArgGluLeuGluArgValLeuAs 193
 Db 576 CTG-----CGCGCTTTCGGGAGCGGAGCGCGCTGCTG-- 609
 Qy 193 nPheProGlnSerAsnLeuCysLeuLysArgGluLysGlnAspLysSerCysSerLeuTh 213
 Db 610 -----ATGGTTAGTTCGTCGGCGAGCGAGCGCGCGCCCTGG 647
 Qy 213 rGlnAlaLeuProSerGluLeuLysValSerAlaAsp---AsnValSerLeuThrGlyAl 232
 Db 648 GTCGTCTCTCCAGCGCGGTGCTCTGGGAAGAGGAGTGAATGTCTGCACAGCAACGC 707
 Qy 232 aValSerLeuAlaSerMetLeuThrGluIlePheLeuLeuGlnAlaGlnGly----- 250
 Db 708 GCTGTGCGCGCGCGGTGCGGCGCGCGGTGTCGCGCGCGCGCGCGGCGCGCGAG 767
 Qy 250 ----- 250
 Db 768 GAGGTTCAAAGAGATCGTCTGAGGAGTGATGGAGTGGTGTGTGTGTAACGCTCGCGCA 827
 Qy 251 ---MetProGluProGlyTrp-----GlyArgIleThrAspSe 262
 Db 828 CTTGTCGCGCGCGCTCCGTCGCTGACCCGCGAGGCGTGGTAGCGAGGATGAAAAAGCT 887
 Qy 262 rHisGln-----TrpAsnThrLeuSerLeuHisAsnAlaGlnPheTy 277
 Db 888 GCACCGCGCGTTCGACGACATGATGAACCGGATCATCGCGAGGAGGCGCGGATCACT 947
 Qy 277 rLeuLeuGlnArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuI1 297
 Db 948 ACTC-----AAACCAACCCACAGTCGTGAGGAAGGTAAAGACTTGTCTGGCTTGTCT 998
 Qy 297 eMetAlaAlaLeuThrProHisProGlnLysGlnAlaTyrGlyValThr----- 314
 Db 999 CTTGCTATGGTG-----CAGGAGCAGGAGTGGCTCGCCCGCGCGAGGA 1043
 Qy 315 -----LeuProThrSerVal-----LeuPheIleAlaGlyHi 325
 Db 1044 CGACAGATCACCGACACGGAATCAAGGCCCTTATCTCTGAATCTATTCTGGCGGCGCAC 1103
 Qy 325 sAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrpThrLeu----- 342
 Db 1104 AGACACAACA-----TCAACCATAGTTGAGTGGACATGAGCAGAGCT 1145
 Qy 343 -----ProGlyGlnProAspAsnThrProProGlyGly-GluLeuValPheGluA 359
 Db 1146 GATTGACACCCACA-TATCTCTCAAGCAGCGCCCAAGAGGAGCTAGATGTTGTTGGGTC 1204
 Qy 359 rGtrpArgArgLeu-----SerAspAsnSerGlnTrpIleGlnValSerLeuValPheG 377

```
Db 407 GCCTCAGCGTCAAGCAGAAATGAGCGCTACGCGCAGTGTGACTCTG-----GTCAAC 457
Qy 113 GluArgThrArgLysThrGlyGluAlaPhe-AlaAlaGlyLeuAlaProAspCysAla11 132
Db 458 TCCACCTCGGCAGACAGAGTGAA---TTCAGCTGCTGGG-----TGCAGCTC 502
Qy 132 eThrValHisThrGlnAlaAspThrSerSerPro-AspProLeuPheAsnProLeuLysT 152
Db 503 TGCAGCGCTACATCTCGCAGGAGCAGCGCAAAACGGGCTC----- 546
Qy 152 hrGlyValCysGlnLeuAspAsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlyG 172
Db 546 ----- 546
Qy 172 LysSerIleAlaAspPheThrGlyHisArgGlnThrAlaPheArgGluLeuGluArgValL 192
Db 547 -----CACCTACATCTCTTTTACAGAGAAAGAGAACTCTTT 583
Qy 192 euAsnPheProGlnSerAsnLeuCysLeuLysArgGluLysGlnAspGluSerCysSerL 212
Db 584 GTACCTTCTCCAGCTACTTCGATGTTGTCTACTTGAACCCGGAACAGACAGCTGTGTG- 642
Qy 212 euThrGlnAlaLeuProSerGluLeuLysValSerAlaAspAsnValSerLeuThrGlyA 232
Db 643 -----TCTTGTGCGGTG 655
Qy 232 laValSerLeuAlaSerMetLeuThrGluIlePheLeuLeuGlnGlnAlaGlnGlyMetP 252
Db 656 ACCGTGCTGTGGCCAAAGT-----CACGCTCCACAGAGGAATC 694
Qy 252 roGluProGlyTrpGlyArgIleThrAspSerHisGlnTrpAsnThrLeuLeuSerLeuH 272
Db 695 CCAGCCAAAGGA-----GATCCAGCCCAATGGAACGGACATGTTTATGAC 739
Qy 272 IsAsnAla-GlnPheTyLeuLeuGln----- 280
Db 740 ATGAAGCGGGCTTGTGTATCTGCACCTCATTTCCGAGCACCAGGTGTGTGTTACTGC 799
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 800 AGGGCGGAGCGCGGGGCGAGATCTCAGATCTCCGTCAAGTAGTACCAGCTGCTCTACGTGGCG 859
Qy 301 LeuThrProHisProProGlnLysGlnAlaTyArgValThrLeuProThrSerValLeu 320
Db 860 GTTCCAGTGGCCCTCCCTCAACA-----ACCATCTTGGCTTCTTCAAAACAAA 907
Qy 321 PheIleAlaGlyHisAsp-----ThrAsnLeuAlaAsnLeuGlyGlyAla 335
Db 908 GTGAAAGTGGGAGCAGCATCAGTGTGCTGTGCACTGTCTGGGGAGCCCGATGTGGAG 967
Qy 336 LeuGluLeuAsnTrpThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeu 355
Db 968 GTGGAGTTCACTGTGATCTTCCAGGCGCAGAGGATGAAGGCT-----GTGACG 1018
Qy 356 ValPheGluArgTrpArg-----ArgLeuSerAspAsnSerGlnTrpIleGlnVal 372
Db 1019 ATCCAAGACACTTGGAGTTGATCCACAGAGGACTGGGACACACCACAGAGAATCTCCAC 1078
Qy 373 SerLeuVal-----PheGlnThrLeuGlnGlnMetArgAspLysThrProLeu 388
Db 1079 AGTGTCATACAGTGAAGACTTCGAGACGATTGAT----- 1114
Qy 389 SerLeuAsnThrProProGlyGluValLysLeuThrLeuAlaGly-----CysGlu 405
Db 1115 -----GCAGGATATTACATTGCACT 1135
Qy 406 GluArgAsnAlaGlnGlyMetCysSerLeuAla 416
Db 1136 GCTCAGATCTTCAAGGACAGACACAGTACGT 1168
RESULT 11
US-10-953-349-34901
```

```
/ Sequence 34901, Application US/10953349
/ Publication No. US20060107345A1
/ GENERAL INFORMATION:
/ APPLICANT: ALEXANDROV, Nikolai et al.
/ TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
/ FILE OF INVENTION: ENCODED THEREBY
/ FILE REFERENCE: 2750-1579PUS2
/ CURRENT APPLICATION NUMBER: US/10/953,349
/ CURRENT FILING DATE: 2004-09-30
/ NUMBER OF SEQ ID NOS: 40252
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 34901
/ LENGTH: 1619
/ TYPE: DNA
/ ORGANISM: Zea mays subsp. mays
US-10-953-349-34901

Alignment Scores:
Pred. No.: 3.68 Length: 1619
Score: 97.00 Matches: 81
Percent Similarity: 36.5% Conservative: 37
Best Local Similarity: 25.1% Mismatches: 107
Query Match: 4.2% Indels: 98
DB: 6 Gaps: 23

US-10-601-319-2 (1-440) x US-10-953-349-34901 (1-1619)
Qy 170 AlaGlyGlySerIleAlaAspPheThr-----GlyHis 180
Db 156 GCGTTGCGCCACATCTCCCTTCACGAGCTCGCAAGCGCATAGCTAGGAGGGCCAC 215
Qy 181 ArgGlnThrAlaPhe-----ArgGluLeuGluArgValLeuAsnPheProGlnSer 197
Db 216 CGGCTCACCTCTTCTCCACGCCAAGAAACACCGCGCGTTAATCCGATCCCTCCAGAA 275
Qy 198 -----AsnLeuCysLeuLysArg-----GluLysGlnAspGlu 208
Db 276 CTCGCGGGGCATATCCGGTGTGTCGATCGCTCGCGCGGTGAGAGCGCTGCCGGAG 335
Qy 209 SerCysSerLeuThrGlnAlaLeuProSerGlu-----LeuLysValSer 223
Db 336 GACTGCGAGGGCGAGCATAGACCTGCCATCCGACGACCTCCGCCCGCTACCTCCGGGTAGCC 395
Qy 224 AlaAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThrGluIlePhe 243
Db 396 TACGAC-----GCCGCTTCGCGCAAAACTCTCGGCCATC--- 431
Qy 244 LeuLeuGlnAlaGlnGlyMetProGlu---ProGlyTrpGlyArgIleThrAspSer 262
Db 432 ---CTGAGAGGAGCGCGG-----CCGAGAGGCGCGACTGGGTCTCATCGACTACGCC 482
Qy 263 HisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyLeuLeuGlnArgThr 282
Db 483 GCGTACTGG-----CGGCTTCGCGCAAAACTCTCGGCCATC---GCG 494
Qy 283 ProGluValAlaArgSerArgAlaThrPro-----LeuLeuAspLeuIleMetAlaAla 300
Db 495 CCGCGCGCGCGCGAGGAGCGCGGTGCGCTTCTCTGAGCGCTTCTCGCGCGCGCG 554
Qy 301 LeuThrProHis-----ProProGln-----LysGlnAlaTyArgValThr 314
Db 555 ACCCTCAGTTTCTATGCGCGCGAGGGGTCTATGGAGCGCGCAAGTACGCCAGGACG 614
Qy 315 LeuProThrSerValLeuPheIleAlaGlyHis-----AspThrAsnLeuAlaAsn 331
Db 615 AAGCGGAGGAGCTCACCGTAGTACCGGCTTACCGTTCGCCACCAACCGTCGGCGAC 674
Qy 332 LeuGly---GlyAlaLeuGluLeuAsnTrpThrLeuProGly---GlnProAspAsnThr 349
Db 675 CGCGGCTTCGAGGCGCGGAGCTA-----TTCAGGACAGCGCTTGTCCCGAGAGCTCC 728
Qy 350 -----ProProGlyGlyGluLeuValPheGluArg 359
-----
```

```
; SEQ ID NO 570
; LENGTH: 14756
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-505-928-570

Alignment Scores:
Pred. No.: 61.3 Length: 14756
Score: 98.50 Matches: 87
Percent Similarity: 34.9% Conservative: 50
Best Local Similarity: 22.1% Mismatches: 140
Query Match: 4.3% Indels: 116
DB: 6 Gaps: 17

US-10-601-319-2 (1-440) x US-10-505-928-570 (1-14756)

Qy 102 SerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGluAla 121
Db 12432 TCCTCTACACTGGCAACAAGCCCTCGTTGTCGACACACACACGTCGCCCATAGAGCA 12373
Qy 122 PheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThrSer 141
Db 12372 TGGTTGGAGAACACGATGACGCTTATCTCAGTGGGTCCCTATGTACAGGCACCT 12313
Qy 142 SerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAlaAsn 161
Db 12312 GCA-----TTTGCAAGTAATAACCTCCAGCAGGTGACGGATTGCACACGCC--- 12268
Qy 162 ValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHisArg 181
Db 12267 -----TCCATTCTGGCAAGGTTGCTGGCGCAGTC-----TTCCGTGGCCGCTCAG 12223
Qy 182 GlnThrAlaPheArgGluLeuGluArgValLeuAsnPhe---ProGlnSerAsnLeuCys 200
Db 12222 GAACGACCTGGAGATACATACCCAGCTCTTCGATGTGTGCATAGTCTCTGGGTTGCT 12163
Qy 201 LeuLysArgGluLys-----GlnAspGluSerCysSerLeuThrGlnAlaLeuPro 217
Db 12162 GTTTAAAGGAGCTCTGCCCATTTCAATAAATGAGTCCATACACCCCTGAAACCAT 12103
Qy 218 SerGluLeuLysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSer 237
Db 12102 ACCAACTTGAGGACTTCTCCATGCTTGTTCCTG-----CTGACGGAT 12058
Qy 238 MetLeuThrGluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTyrGly 257
Db 12057 GTGCCCAACCAAAACACATAGTTATCCAGGTTTCCAGGTTTCCAGAGTCCC---TGGGGC 12001
Qy 258 -----ArgIleThrAspSerHisGln 264
Db 12000 TGTGCCCGATGCAGTATGACTGGTCTAGACCAACGAGCATAGTTTCCATTCACCTC 11941
Qy 265 TrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGlnArgThrProGlu 284
Db 11940 CAGGGCCACCTGCGTGCCTGCCCATCTATTGACCTGAATGCTCTGAACAGAGAC----- 11887
Qy 285 ValAlaArgSerArgAlaThrPro----- 292
Db 11886 -----AATTCAGGGCCACTTCCACAGTCAAACTGTACTGCGAGCCTCCATGATGAAT 11833
Qy 293 LeuLeuAspLeuIleMetAlaAlaLeuThrProHisProGlnLysGlnAla-TyrGln 312
Db 11832 CTCAAGATGCTATAGTACTTCTCGAGCATACGACCAACCGCATCGCTGGGAATATGT 11773
Qy 312 yValThrLeu----- 315
Db 11772 TCTGAGCCTCATGTCAGTTTTCATCTCTAATTGTTTTCATTTTCCGTCAGACGGTATTT 11713
Qy 316 -----ProThrSerValLeu---PheIleAlaGlyHisAspThrAsnLeuAl 330
Db 11712 CAGTAGCTGTTCCAGTGTGATAGATGAACTCCCTGGGCACTGACCAACCTGCC 11653
Qy 330 aAsnLeuGlyGlyAlaLeuGluLeuAsnTrpThrLeuProGlyGlnProAspAsnThrPr 350
```

; APPLICANT: Bio-oriented Technology Research Advancement Institution.									
; APPLICANT: The Institute of Physical and Chemical Research.									
; APPLICANT: Foundation for Advancement of International Science.									
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF									
; FILE REFERENCE: MOA-A0205V1-US									
; CURRENT APPLICATION NUMBER: US/10/449,902									
; CURRENT FILING DATE: 2003-05-29									
; PRIOR APPLICATION NUMBER: JP 2002-203269									
; PRIOR FILING DATE: 2002-05-30									
; PRIOR APPLICATION NUMBER: JP 2002-383870									
; PRIOR FILING DATE: 2002-12-11									
; NUMBER OF SEQ ID NOS: 56791									
; SOFTWARE: PatentIn Ver. 2.1									
; SEQ ID NO 12801									
; LENGTH: 3287									
; TYPE: DNA									
; ORGANISM: Oryza sativa									
; PUBLICATION INFORMATION:									
; DATABASE ACCESSION NUMBER: AK110149									
; DATABASE ENTRY DATE: 2001-12-06									
US-10-449-902-12801									
Alignment Scores:									
Pred. No.:		6.51		Length:		3287			
Score:		99.00		Matches:		104			
Percent Similarity:		32.7%		Conservative:		56			
Best Local Similarity:		21.2%		Mismatches:		174			
Query Match:		4.3%		Indels:		156			
DB:		6		Gaps:		20			
US-10-601-319-2 (1-440) x US-10-449-902-12801 (1-3287)									
Qy	1	MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla	20						
Db	334	CTGCAAGCGCACTCCCGCTTCATCTCCCGAGCATCCCGCCATCCGACCTCA	390						
Qy	21	PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly	40						
Db	391	-----AAAGGSCCTCCGTCTCTGCCATCTTCAGTCAGAGCGACTCGATCAACA	441						
Qy	41	Val--ArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpP	60						
Db	442	GTCTCTCGAGCACCCACCCGGTTCACGCCACAGAGCAGCATGCCATCCCAGAAGCGCAAC	501						
Qy	60	roThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyL	80						
Db	502	CA-----CTACCAACAACACAGGAGCTGCAGGCTCAAG	534						
Qy	80	euglyHisTyxGlnArgGln-----	88						
Db	535	TCTCGACCTCACCAACACAGGTGACGGTCTCAACGAGAGCTCGTCAATCCTTCAATC	594						
Qy	88	euValAlaAsp-----	91						
Db	595	GCATCGCGCAGCTCGAGGACGACTACTCGGACGCCAAGAGACCATTTCGACCATGTCGG	654						
Qy	92	-----	93						
Db	655	CAAAAGATCAAGAGGCTCGAAAGAGGCGCGCAGAGCACTCGACGCCCTCAACACTGGGC	714						
Qy	93	euLeuAlaLysLysGlyCysProGlnSerGlyGlnValAlaIleIleAlaAspValAspG	113						
Db	715	TCCTCGTGAAGAGGCTCAGCTCAGCAGGAGATGCAGAGGATGATGGCCGTGTCATCG	774						
Qy	113	luArGThrArgLysThrGlyGluAlaPheAlaGlyLeuAlaProAspCysAlaIleT	133						
Db	775	AAGAGACCGCCCAACCGCGCAAGGCA-----	800						
Qy	133	hrValHisThrGlnAlaAspThrSerSerProAspProLeuPheAsnProLeuLysThrG	153						
Db	801	-----GAGATGACCAAGACCAAGATCAGGCGCGAGCTCGACGAGCTCAGCTCGA	849						
Qy	153	lyValCysGlnLeuAspAsnAlaAenValThrAspAlaIleLeuSerArqAlaGlyVLs	173						

```

QY 273 enAlaGlnPhe-----TyrLeuLeuGlnA 281
    ::::
Db 713 CATCCACAAAGAGGCTGATCCCTGCTCCACTCCACAGGCCCGACACTGCACTCCC 654
QY 281 rgThrProGluValAlaAAsrArgAlaThrProLeuLeuAspLeuIleMetAlaAlaL 301
    ::::
Db 653 AGACCCCTGGGGTGGCTTGGCCGGCAGCCCTTGCAGCCAGGCTCCAGGCGGAAG 594
QY 301 eutThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeuP 321
    ::::
Db 593 CAGAGCT-CATCCA-----GGCGTACTGCCCCACGACGCGCCCGGT 553
QY 321 heileAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp 341
    ::::
Db 552 ATCTGGAACACACTGCCTGCTG----- 527
QY 341 hrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrpA 361
    ::::
Db 526 -----CCCGTGGGAGGAGGAGCGGCATTCTCTGCTGATT 493
QY 361 rgArg-LeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPhe 376
    ::::
Db 492 CCAGGTCTGTGCCACATGTCATCTGTGTGTCAGGGAACCTCGAGTTC 445

RESULT 7
US-11-217-529-274
; Sequence 274, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 274
; LENGTH: 2259
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
; US-11-217-529-274

Alignment Scores:
Pred. No.: 2.47 Length: 2259
Score: 101.00 Matches: 93
Percent Similarity: 35.3% Conservative: 45
Best Local Similarity: 23.8% Mismatches: 138
Query Match: 4.4% Indels: 116
DB: 7 Gaps: 21

US-10-601-319-2 (1-440) x US-11-217-529-274 (1-2259)
QY 101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
    ::::
Db 7 CAGAGCAACAAGATCGGTAGCTGCTGACTCCGAC-----ACGACAGGAGCGCGCGAG 60
QY 121 AlaPheAlaGlyLeuAlaProAsp-----CysAlaIleThrValHisThrGln 137
    ::::
Db 61 TCGCGCTCCGCTCGCTCTTCGACAGAAGCAGCTGCCAAGCTGCTCGTCCACACG--- 117
QY 138 AlaAspThrSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeu 157
    ::::
Db 118 -----CTCAAG-----GAATTG 129
QY 158 AspAsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPhe 177

```

RESULT 8

```

US-10-449-902-12801
; Sequence 12801, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.

```

```

Db 130 GGCTACGACTCCCGCGGACACGAGTCTCTCGAGAGCGCGGG---TACCAGAACGAG 186
    ::::
QY 178 ThrGlyHisArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSer 197
    ::::
Db 187 TCGAACCATATCCAGACCTTCTTC-----AGGCTCTCAAGACCGCCAGTTC 234
QY 198 AsnLeuCysLeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuPro 217
    ::::
Db 235 GACCTCATGACTGGCGC-----AGCGTGTTCGCTGCCCTCGCACATAGC 282
QY 218 SerGluLeuLysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSer 237
    ::::
Db 283 TCGCCGCTCAGGCGCTGCTGCTGCAACGCTCTCTCGCCCGCAGCGCCACCGAGCAGC 342
QY 238 MetLeuThrGluIlePhe-----LeuLeuGlnGlnAlaGlnGly-----Met 251
    ::::
Db 343 GCCACCGGCTGCTGTTTGACCACTGCTTCAGCAGCTGCGAGGACTGCGAGCTGCTGATG 402
QY 252 ProGluProGlyTrpGlyArgIleThrAspSer-----HisGlnTrpAsn 266
    ::::
Db 403 GCCTCTGTGGCGCGCTCCACTTCTCGGACGCGCAGGTCCACAGCTCAGGAACACTAGTC 462
QY 267 ThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGlnArgThrProGluValAla 286
    ::::
Db 463 GAGATCATGATCTCTGTCACAGCGAGATCTCTCTCGAGTCTTCCACCCGCGCTGCAGCT 522
QY 287 ArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaLeuThrProHis----- 304
    ::::
Db 523 GCTGCT-----GCTGCCGCGCGCCCGCCACCGCGGC 552
QY 305 ProGlnLysGlnAlaTyrGlyValThrLeuProThr----- 317
    ::::
Db 553 CCGCCC-----ACCACGCTCGCGCTCTGTACTGCGCAAGGTCTCTC 594
QY 318 -----SerValLeuPheIleAlaGly-----HisAsp 326
    ::::
Db 595 AAGTGTCTCATCGAGATCTGGGACGCGCTGTGTGTCCTCAATGACCACTTCTCAACGAG 654
QY 327 ThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrpThrLeuProGlyGlnPro 346
    ::::
Db 655 GAGAACATCTTCAACCC-GGAAACAACCTTGAGAGAGCTCTCCACAGTACTCTACGAACCC 713
QY 347 Asp-----AsnThrProGlyGlyGluLeuValPhe 357
    ::::
Db 714 GAAGCTCAGCGCGCACTTGAACTCGAGCGCGACCACTGATCGAGCCCATCTCCAATA 773
QY 358 GluArgTrpArg-----LeuSerAspAsnSerGlnTrpIle 370
    ::::
Db 774 CATCGACCGCGAACGAGCTCGTCCCAAGGCGCGCTCTTGCACCTCTTGAAGCAGGCCAT 833
QY 371 GlnValSerLeuValPheGlnThrLeuGlnMetArgAsp-----LysThrProLeu 388
    ::::
Db 834 CAAGTACCAAGCAGCTCCAGACATCTTCAACATCTCCGACCCCGCAGCAGACGCCCTTT 893
QY 389 SerLeuAsnThrProGlyGluValLysLeuThrLeuAlaGlyCysGluGluArgAsn 408
    ::::
Db 894 CTCCTCGAGTCCACCGCATCAA-----TCTGCTGCAGGA-----CAACATCTC 938
QY 409 AlaGlnGlyMetCysSerLeuAlaGlyPheThrGlnIleValAsnGluAlaArgIlePro 428
    ::::
Db 939 GCACGACTTGACTGTGACCTTCCAGGA----- 965
QY 429 AlaCysSerLeuArgSerHisHisHisHis 439
    ::::
Db 966 -----CTGAAGACCATCCAGACACCCAC 989

```

```

QY 178 -----ThrGlyHisArgGlnThrAlaPheArgGluLeu 188
Db 4168 CATGGGTCCAGCGCTTACAGTGGGACAGCCCATGTCAGGTGGCCGCGCAACGCCCTG 4227

QY 189 Glu-----ArgValLeuAsnPheProGlnSerAen-----LeuCysLeuLysArg 203
Db 4228 CAGCTGGGGTGAGATGTGCCTCAGTGACAGGACACCTACGTCTGCATGACCGC 4287

QY 204 GluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeuLysValSer 223
Db 4288 AACTCTCTGGGCTCAGTCAGCACCATGGGAGCTGCAGCCAGAGAGGTGTGCACGTGTA 4347

QY 224 AlaAsp-----AsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 4348 GCTGAGCAGGGCTGCTGCTGAGGCACAGCGCTGAACCTGAGCTGTGCCTCCCT 4407

QY 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGly----- 255
Db 4408 AGT-----GGCCCTGGGCACATAGGCACTCC 4434

QY 256 -----Trp-----GlyArgIleThrAspSerHisGlnTrpAsnThrLeuLeu 269
Db 4435 ACCTTGTCTGTTCGGAACGGTCGGCAGCTACACACAGAGTCTGTGCCACCTTACC 4494

QY 270 SerLeuHisAsnAlaGln-----PheTyrLeuLeuGlnArgThrProGlu 284
Db 4495 TTCACCCATGTGGCCGCGCCCAAGCTGGCTTGTACCACTGCCAGCTGAGCTCCCGCC 4554

QY 285 ValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAlaLeuThrProHis 304
Db 4555 GGGGCTGCCACCTCTGCTCCAGTCTTGTTCGGGTGCTC-----TAC 4596

QY 305 ProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeuPheIleAlaGly 324
Db 4597 CCTCCCAAG-----ACGCCCAACCATGACGTGTTTTTGTGGAGGCC 4635

QY 325 HisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlu-LeuAsnTrpThrLeu----P 343
Db 4636 GAG-----GGTGGCATCCAGGGCATTTCTGGACTGCCAGTGGACAGTGAGGCC 4683

QY 343 roGlyGlnProAspAsnThrProPro-GlyGlyGluLeuValPheGluArgTrpArgArg 362
Db 4684 CTAGCCAGCCTGAC---CCTCCACTGGGAGTGGCTGGTGGCTCCAGCCAGCCTCAG 4740

QY 363 LeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGlnGlnMet 382
Db 4741 GCTGCCCTGCCAGCCGACATCCGCTCTCA----- 4773

QY 383 ArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThrLeuAla 402
Db 4774 -----GCCATCCCAATGGCTTGGAGTGACATGGAGAGCTGAAGCCAGTGACCCAG 4827

QY 403 Gly-----CysGluGluArgAsnAlaGlnGlyMetCysSerLeuAla----- 416
Db 4828 GGGGAGTATGTGTCTGGCTTCCATGCTGGCTGTGGCTCTGTGCTGTGCCACTACTTTC 4887

QY 417 GlyPheThrGlnIleValAsnGluAlaAlaArgIleProAlaCysSerLeu 432
Db 4888 GGAACACAGCCCTGCA-TCGGCTGCTATCTGTTCAGCACCTTCTCTG 4934

```

RESULT 6

```

US-11-293-697-1957/c
; Sequence 1957, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28

```

```

; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1957
; LENGTH: 1945
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-1957

```

```

Alignment Scores:
Pred. No.: 1.79 Length: 1945
Score: 101.50 Matches: 95
Percent Similarity: 34.7% Conservative: 43
Best Local Similarity: 23.9% Mismatches: 148
Query Match: 4.4% Indels: 114
Db: 7 Gaps: 13

```

US-10-601-319-2 (1-440) x US-11-293-697-1957 (1-1945)

```

QY 5 LeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAlaPheAlaGlnSer 24
Db 1372 CTGATTCGGTT-----CTAGTGGTGCACAGACGGCCCGCTCAGCT----- 1331

QY 25 GluProGluLeuLysLeuGluSerValValIleValSerArgHisGlyValArgAlaPro 44
Db 1330 -----CTCAGAGTTCACCTCAATCTTCTCAGCTTGGACCTACGGG----- 1292

QY 45 ThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpProThrTrpProVal 64
Db 1292 ----- 1292

QY 65 LysLeuGlyTrpLeuThrProArgGlyGlyLeuIleAla----- 78
Db 1291 AAATTTGGGGTTTTCACAGAACTGACGGGCAATCTGGATGTTCACTGAAGAGCAGCTAG 1232

QY 79 -----TyrLeuGlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeu 93
Db 1231 GTGATTCACGGGTAACTATTGGAGCTACACAAATACACGCGTT-----CTC 1184

QY 94 LeuAlaLysLysGlyCysProGlnSerGlyGlnValAlaIleIleAlaAspValAspGlu 113
Db 1183 AAATCACCTAAAGGCTCTTTCAAATAAGGATGTGTACTCTCCAAAGAC----- 1133

QY 114 ArgThrArgLysThrGlyGluAlaPheAlaAlaGlyLeuAlaProAspCys-AlaIleTh 133
Db 1132 -----TCCTGTCTCTGGATGTGCCAAAC 1109

QY 133 rValHisThrGlnAlaAspThrSerSerProAspProLeuPheAsnProLeuLysThrGl 153
Db 1108 A-----AAGGATGCTCAGACTCTGACGCTGGAGCTACAGAGCACCTTGGCCCTCGG 1055

QY 153 YValCysGlnLeuAspAsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlyGlySe 173
Db 1054 GTCCGAGAAAGACCCAGCACCAAGACGTGCCCAAGCTTCCAGCTTCGCGCTCAGCTGT-GA 996

QY 173 rIleAlaAspPheThrGlyHisArgGlnThrAlaPheArgGluLeuGluArgValLeuAs 193
Db 995 TGTGCCACGCGGGGACAGGACAC-----AGCCCATGTGGCGGCAGCTGCCAAGAGCTGCT 939

QY 193 nPheProGlnSerAsnLeuCysLysLeuLysArgGluLysGlnAspGluSerCysSerLeuTh 213
Db 938 TCAGCCCAAGTAAACACACAGA---GTTGCCTCACAGCCCTCTCTCTGTCTGTGCTGCAA 882

QY 213 rGlnAlaLeuProSerGluLeuLysValSerAlaAspAsnValSerLeuThrGlyAlaVa 233
Db 881 CGCAGCTTGTGGACAGGGAGCTCCGCTTGCAGCTGACTCTCTGGAGGCGCTGGGCCAGCT 822

QY 233 lSerLeuAlaSerMetLeuThrGluIlePheLeuLeuGlnGlnAlaGlnGlyMet-ProG 253
Db 821 GGACCTCAGAGCCCAAAAGTGTCTCCAAAT-----CCTGCCCGCCAGCTGTGTCAG 774

QY 253 luProGlyTrpGlyArgIleThrAspSerHisGlnTrpAsnThrLeuLeuSerLeuHisA 273
Db 773 AACCGAAACCTGGAGGCTTGACAGCTTCTTACCAGAGAGTTTCCACGCTGCTCTCCATTCT 714

```

```
QY 26 ProGluLeuLysLeuGluSerValValIleValSerArgHisGlyValArgAlaProThr 45
Db 239 CGGAAGTGTCCGCCAGCGCATGCTAGACGTGGCCACCAGGGTGTAAAGCCCTGGCAC 298
QY 46 LysAlaThrGlnLeu-----MetGlnAspValThrProAspAlaTyrP 59
Db 299 ATCTCTGCCAGCTCCCGCTGAGCGCATGCTGCTGCTGAGGCGCTGT---GACAAATGACACTGTG 358
QY 60 ProThrTrpProValLysLeuGlyTyrLeu-----ThrPro 71
Db 359 GAGCTGGCGACATCCCGCTGGAGTGTAGTGGGGCTCCAGCCCAAGGTGGCCACCC 418
QY 72 ArgGlyGlyLeuLeuAlaTyrLeuGlyHisTyrGlnArgGlnArg----- 87
Db 419 AAG-----GTGGTGGAGAGATTGGGACTACAAACGCCAGCAACCTTACCATGTT 469
QY 88 -----LeuValAlaAspGlyLeuLeuAlaLysLysGlyCysProGlnSerGlyGlnVal 105
Db 470 GCCTGGGAGATCCGAGACCGGCTCTGCTGCTGAGGCGCTGT---GACAAATGACACTGTG 526
QY 106 AlaIleIleAlaAspValAspGluArgThrArg---LysThrGlyGluAlaPheAlaAla 124
Db 527 CCCAGTGTAGCTCCATTATAGATCATCCGACCAAGTGCAGCAACCATTC----- 580
QY 125 GlyLeuAlaProAspCysAlaIleThrVal-----HisThrGlnAla 138
Db 581 AACCTCCCTATGGACAGCTGGTGGCCACCAGTCCCTGAGTCCCGGACACAGCTGATC 640
QY 139 AspThrSerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAsp 158
Db 641 CCCAGCTCAGCTGTAACCTCCCGGAGTCACCC-----CAGTCGGAT 682
QY 159 AsnAlaAsnValThrAspAlaIle-Leu---SerArgAlaGlyGlySerIleAlaAspPh 177
Db 683 TCCCTGGGCTCCACTACTCATCATATAGTGGCTCTGCGCATCGCTGAGCTGGCAGCGAC 742
QY 177 eThrGly-----HisArgG1 182
Db 743 AAGAGGAATGGTACAGTATGATGATAGTCCGACCTAAGCATTTGACTCACAGAGC 802
QY 182 nThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCysLeu 202
Db 803 AGCAGCAGCGGACCCCGAAGACACCTTCGACAGCATGCTTCAGCCAGCACCCACTCGAG 862
QY 202 sArgGluLysGlnAspGluSerCysSerLeuThrGln---AlaLeuProSerGluLeu 221
Db 863 CCCTCGAGTGCCCATTTGAGCGGACGACACTACCCAGAGCGCTATGCTCCCGCCAGCCAC 922
QY 221 sValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThrG1 241
Db 923 ACCAAGCGGACGAGCGCTCTACCCGCTGCTGCTCAACAGCACCTTGGAGCAGCGG 982
QY 241 uilePheLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrPglyArgIleThrAs 261
Db 983 A-----AGGCCACCTGACCCCTTCCAAACAGCGCACTGGGGCGCAACCTCTCG 1030
QY 261 pSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGlnAr 281
Db 1031 -ACTCACCACTTACCCCTGGTGGGAGATTCCTCACTACCTTGGCCATAAGCAGGA 1089
QY 281 gThrProGluValAlaArgSerArgAlaThrPro-----LeuLe 294
Db 1090 AACCCCGAGGTGTCCAGTCTTAGCTCCACCCCTGCTTTTATCTAGTCCCGCCCTTTT 1149
QY 294 uAspLeuIleMetAlaAlaLeuThrProHisProGlnLysGlnAlaTyrGlyValTh 314
Db 1150 GGATCTG-----CAGCAAGTCGCTCCGG 1173
QY 314 rLeuProThrSerValLeuPhe-----IleAlaGlyHisAspThrAnLe 329
Db 1174 GGTCCCGCCCTTCAATGCCTTTCCCATGTCGCTCGTGTACGGGCGAGTTCACGGGCCA 1233
```

```
QY 329 uAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrpThrLeuProGlyGlnProAspAnTh 349
Db 1234 GGCCCTCTCTCAGGCGGAGAGATGTGGGCCACGCTGCCCGGATACCCACCCACAT 1293
QY 349 rProProGlyGlyGlu 354
Db 1294 CCCACACGCGGACAG 1309
RESULT 5
US-11-270-040-5
; Sequence 5, Application US/11270040
; Publication No. US20060110762A1
; GENERAL INFORMATION:
; APPLICANT: Kapil, Sanjay
; APPLICANT: Kim, Jeong-Ki
; TITLE OF INVENTION: PORCINE REPRODUCTIVE AND RESPIRATORY
; SYNDROME VIRUS RECEPTOR COMPONENTS AND USES THEREOF
; FILE REFERENCE: 14337/004001
; CURRENT APPLICATION NUMBER: US/11/270,040
; CURRENT FILING DATE: 2005-11-09
; PRIOR APPLICATION NUMBER: 60/680,297
; PRIOR FILING DATE: 2005-05-12
; PRIOR APPLICATION NUMBER: 60/626,788
; PRIOR FILING DATE: 2004-11-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 5193
; TYPE: DNA
; ORGANISM: Sus scrofa
US-11-270-040-5
Alignment Scores:
Pred. No.: 1-27 Length: 5193
Score: 109.50 Matches: 105
Percent Similarity: 35.2% Conservative: 56
Best Local Similarity: 22.9% Mismatches: 190
Query Match: 4.8% Indels: 108
DB: 7 Gaps: 21
US-10-601-319-2 (1-440) x US-11-270-040-5 (1-5193)
QY 22 AlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGlyVal 41
Db 3748 GCCAACACATCCCTGGAGCTGCGCTAGAGGGCGTGCAGTGGCAGCTCCATCGGCC 3807
QY 42 ArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTyrProThr 61
Db 3808 ACTGTCCGAGGGGGCCCTGTCAAGTACCTGTGAAGACCTGTGTCGCCGCCA--- 3864
QY 62 TrpProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeuGly 81
Db 3865 ---CCCACTCTCTATGCTGG----- 3882
QY 82 HisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysProGln 101
Db 3883 ---TACCACACACAGCGTGGTGGCTGAGAGGGGTGCGCTGCTCTCTCTGTTTCCA--- 3936
QY 102 SerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGluAla 121
Db 3937 -----GCGGCTACACGGGCTCACGGGGCGCTATACCTGCCAGGTCCAGGATGCC 3987
QY 122 PheAlaAlaGlyLeuAlaProAspCysAlaIleThrVal---HisThrGlnAlaAsp--- 139
Db 3988 CAGGGCAGCAGCATCTCCAGCCCGCAGCACCTCATCTCTATGCCCTCGGGAGTCT 4047
QY 140 -----ThrSerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeu 157
Db 4048 GTCTTTTCTCTCTCTGGGACTCAAGGGCCAGCCCTATGGCCGTGTACAGTCACTGTG 4107
QY 158 AspAsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPhe 177
Db 4108 GACAGCGAGCCACCTCCCGAGATGACCTGTCTCCCATGATGGCAAGGTGTGTGGCCACGAGC 4167
```

QY	187	uLeuGluArgValLeuAsnPheProGlnSerAsnLeu----	CysLeuLysArgGluLysGlu	206
DB	244	CTTGGCCGCGCCACACAGCTTCATCGCTCCGTGATCTTCTCGTGTGCGCAC-CGAGT	186	
QY	206	nAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeuLysValSerAlaAspAs	226	
DB	185	CGATCGGCTCGCAGCTCACGAGGACACCGACACCTGCAGGCGCGCTGCTGC	126	
QY	226	nValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThrGluIlePheLeuLeuGlu	246	
DB	125	CCAGCTAGGACGGGCGCGCGC-----	101	
QY	246	nGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThrAspSerHisGlnTrp	265	
DB	100	-GGTGCCAGGCGGAGGAGGACCGGCGGTGCAGAGGTGCAGACGACATGGCTGG	44	

RESULT 3

US-10-449-902-7723/c

; Sequence 7723, Application US/10449902

; Publication No. US20060123505A1

; GENERAL INFORMATION:

; APPLICANT: National Institute of Agrobiological Sciences.

; APPLICANT: Bio-oriented Technology Research Advancement Institution.

; APPLICANT: The Institute of Physical and Chemical Research.

; APPLICANT: Foundation for Advancement of International Science.

; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF

; FILE REFERENCE: MOA-A0205V1-US

; CURRENT APPLICATION NUMBER: US/10/449,902

; CURRENT FILING DATE: 2003-05-29

; PRIOR APPLICATION NUMBER: JP 2002-203269

; PRIOR FILING DATE: 2002-05-30

; PRIOR APPLICATION NUMBER: JP 2002-383870

; PRIOR FILING DATE: 2002-12-11

; NUMBER OF SEQ ID NOS: 56791

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 7723

; LENGTH: 2822

; TYPE: DNA

; ORGANISM: Oryza sativa

; PUBLICATION INFORMATION:

; DATABASE ACCESSION NUMBER: AK105071

; DATABASE ENTRY DATE: 2002-08-28

US-10-449-902-7723

Alignment Scores:		
Pred. No.:	0.18	Length: 2822
Score:	114.50	Matches: 83
Percent Similarity:	37.4%	Conservative: 30
Best Local Similarity:	27.5%	Mismatches: 108
Query Match:	5.0%	Indels: 82
DB:	6	Gaps: 16

US-10-601-319-2 (1-440) x US-10-449-902-7723 (1-2822)

QY	12	LeuIleProLeuThrProGlnSerAlaPheAlaGlnSerGluProGluLeuLysLeuGlu	31
DB	2306	CTTCTCCGCTCCAGGTCAGCTTCCTTCCTTGGTGTGAGCCGGA-----CAG	2256
QY	32	SerValVal-LleValSerArgHisGlyValArgAlaProThrLysAlaThrGlnLeuMe	51
DB	2255	CTCGTCTCGTGGGTCCAGCAC-----ACGCCGTTGAT	2226
QY	51	tGlnAspValThr-ProAspAlaTrpProThrTrpProValLysLeuGly-----	67
DB	2219	GTCCAGATCTCCCCCT-----TGGCGCATCGGAGCAGCGCTCGATGATGGGTC	2166
QY	68	-----TrpLeuThrProArgGlyGly-GluLeuIleAlaTyr	79
DB	2165	GGAGCGGACGAGGAGCAACTGG-----AGTAGAGGAGCTCCACTTGTCCACGC	2112
QY	80	LeuGlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGly----	98

```

Query Match:      5.1%      Indels:      78
DB:                6        Gaps:        14

US-10-601-319-2 (1-440) x US-10-449-902-14217 (1-1482)

QY      19 SerAlaPheAlaGlnSerGluProGluLeuLysLeuGlnSerValAlaIleValSerArg 38
      1957 TCGATCTTACACCTTCTCCGCGCCACCGGTGAGCTTCCCTTCTGTGTGTGTAGCCGGAC 898
QY      39 HisGlyValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThr-ProAspAl 58
      897 AGCTGCTCTGCGTGGCTCCACGACGCGCTTGAATGTGACAGATCTCCCTC----- 843
QY      58 ATTPProThrTTPProValLysLeuGly----- 67
      842 -TGGCGGACGATGGGAGACGCGTCTGATGATGGGGTTCGACGACGAGGACGAGACGAAAC 784
QY      68 -TTPLeuThrProArgGlyGly-GluLeuIleAlaTyrLeuGlyHisTyrGlnArgGlna 87
      783 TTGG-----AGTAAAGAGCTCCACTTGTCCACCGCTCGACACGACGAGCTAG 730
QY      87 rgleuValAlaAspGlyLeuLeuAlaLysGly-----CysP 100
      729 ACGAGGTGCGAGATGAGACTGGAGATCTTCAAGGTGGGATGCGCTTCACTCCAGCGTC 670
QY      100 roGlnSerGlyGln-----ValAlaIleIleAlaAspVala 112
      669 CGCTCGGTGGGATGAAAGGCGCGGATGAAATGAGCGTTCCTTCTGCGACGCTG 610
QY      112 spGluArgThrArgLysThrGly-----GluAlaPheAlaAlaGlyLeuAla-----P 128
      609 ACGAGCGTGACTCAGGCGCGAGCTGCTTGAAGCTCTGATGAGCGGTCTGCGCTTCTTG 550
QY      128 roAspCysAlaIleThrValHisThrGlnAlaAspThrSerSerProAsp-----P 145
      549 AGGACGTGTGTGTGAAGCTCGCGAGAGCGCGCTGCGGTAGAGACGACGAGCGG 490
QY      145 roLeuPheAsnProLeuLysThrGly-----ValCysGlnLeuAsp--A 159
      489 ACCTTCTTGAACGCGGATGCGGTGAGGGGAGGTGATGATCTCTCGCTGGAATCTCC 430
QY      159 snAlaAsnValThrAspAlaIleLeuSerArg-----AlaGlyLysSerIleA 175
      429 TGGTTCATGTGTGAAGACCTCCACGACGCTCCGAAAGGGCGGAGAGACACACC 370
QY      175 laaAspPhe-----ThrGlyHisArgGlnThrAlaPheArgGlnLeuG 189
      369 GCCTCTGCGCGCGCGCGACCTTGGCGCGCGCACACGCTCATCGCTCGTATCTTC 310
QY      189 luaIrgValLeuAsnPheProGlnSerAsnLeuCysLeuLysArgGlnLysGlnAspGln 209
      309 TGCCTGTGTGCGCAC-----CGAGTCCAGTGGCG 283
QY      209 erCysSerLeuThrGlnAlaLeuProSerGluLeuLysValSerAlaAsnValSerL 229
      282 CTCCGCACTCAACGACGAGACGACGACGACCTGACGCGCGCTCTCTCGCGAGCTA 223
QY      229 eurtHrGlyAlaValSerLeuAlaSerMetLeuThrGlnIlePheLeuLeuGlnAlaG 249
      222 GGAGCGGAGCGCGCGG-----GGTGGCC 199
QY      249 lngIlyMetProGluProGlyTTPGlyArgIleThrAspSerHisGlnTTP 265
      198 AGGGCGGAGAGACCAAGCGGTGAGAGTGGCAGACGACATGGCTGG 149

```

```

APPLICANT: Foundation for Advancement of International Science.
TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
FILE REFERENCE: MOA-A0205Y1-US
CURRENT APPLICATION NUMBER: US/10/449,902
CURRENT FILING DATE: 2003-05-29
PRIOR APPLICATION NUMBER: JP 2002-203269
PRIOR FILING DATE: 2002-05-30
PRIOR APPLICATION NUMBER: JP 2002-383870
PRIOR FILING DATE: 2002-12-11
NUMBER OF SEQ ID NOS: 56791
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 21879
LENGTH: 1470
TYPE: DNA
ORGANISM: Oryza sativa
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: AK072104
DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-21879

Alignment Scores:
Pred. No.: 0.0712      Length: 1470
Score: 114.50      Matches: 83
Percent Similarity: 37.4%      Conservative: 30
Best Local Similarity: 27.5%      Mismatches: 108
Query Match: 5.0%      Indels: 82
DB: 6      Gaps: 16

US-10-601-319-2 (1-440) x US-10-449-902-21879 (1-1470)

QY      12 LeuIleProLeuThrProGlnSerAlaPheAlaGlnSerGluProGluLeuLysLeuGln 31
      841 CTCTCCCGCTCCACGCGTCACTTCTCTGTGTGTGTAGCCGGA-----CAG 791
QY      32 SerValVal-IleValSerArgHisGlyValArgAlaProThrLysAlaThrGlnLeu 51
      790 CTGCTCTCGGTGGCGCTTCACGAC-----ACGCCGTGAT 755
QY      51 GlnAspValThr-ProAspAlaTTPProThrTTPProValLysLeuGly----- 67
      754 GTGCGCATGCTCCCT-----TGGCGCATGGGAGAGCGCTGATGATGGGCTC 701
QY      68 -----TTPLeuThrProArgGlyGly-GluLeuIleAlaTyr 79
      700 GGAGCGGACGAGGAGACGAACTGG-----AGTAAAGAGCTCCACCTTGTCCACGCG 647
QY      80 LeuGlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGly--- 98
      646 CTGCGACACGAAACGAGATGACGAGTGGAGTGGATGGATCGGAGTCTTCAACGGTGG 587
QY      99 -----CysProGlnSerGlyGln----- 104
      586 GATGCCGTTCACTCCAGCGTCCGCTGCGGGATGAAGGGCGCGGATGAAGTAGGC 527
QY      105 ValAlaIleIleAlaAspValaAspGluArgThrArgLysThrGly-----GluAlaPhe 122
      526 GTTCCCTTCTTGTGCGACGCTGACGAGGTGACTGACGCGGACGCTGTTGAGCTCTC 467
QY      123 AlaAlaGlyLeuAla-----ProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
      466 GATGCGGCTCTCGGCTTCTTGAAGACGTTGTTGTAACCTGCGGAGAGGCGCGCTC 407
QY      141 SerSerProAsp-----ProLeuPheAsnProLeuLysThrGly----- 153
      406 GCCGCTGAGGACGACGAGGCGGACCTTCTTGAACGCGGCGGATGAGGGGAGAGTC 347
QY      154 ---ValCysGlnLeuAsp---AsnAlaAsnValThrAspAlaIleLeuSerArg----- 169
      346 GATGCTCTCGCTGTGGAATCTCGTGTGTTGATGAGACCTCCACGACGCGCTTCCGA 287
QY      170 -----AlaGlyLysSerIleAlaAspPheThrGlyHisArg-GlnThrAlaPheArgGln 187
      286 GAAGGCGGAGGAGAC-----CACGCGCTCTCGCGCGCGCGCAC 245

```

```

RESULT 2
US-10-449-902-21879/c
; Sequence 21879, Application US/10449902
; Publication No. US2006012305A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.

```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 14, 2006, 14:22:06 ; Search time 33.8073 Seconds
(without alignments)
4595.351 Million cell updates/sec

Title: US-10-601-319-2
Perfect score: 2302
Sequence: 1 MKALIFPLSLIPLTPSA.....TVNBARIPACSLRSHHHHH 440

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 290227 seqs, 117694381 residues
Total number of hits satisfying chosen parameters: 580454

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+.p2n.model -DEV=xlh
-O=/abs/ABSSWEB_spool/US10601319/runat_12062006_180105_27789/app_query.fasta_1
-DB=Published Applications NA New -QFMT=fastap -SUFFIX=rnpbn -MINMATCH=0.1
-LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSTR=500 -MINLEN=0
-MAXLEN=200000000 -HOST=abse03h
-USER=US10601319@CGN_1_1_28 @runat_12062006_180105_27789 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DSPBLCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New:*
1: /EMC_Celestra_SIDS3/ptodata/2/pubpna/US09_NEW_PUB_seq.*
2: /EMC_Celestra_SIDS3/ptodata/2/pubpna/US06_NEW_PUB_seq.*
3: /EMC_Celestra_SIDS3/ptodata/2/pubpna/US07_NEW_PUB_seq.*
4: /EMC_Celestra_SIDS3/ptodata/2/pubpna/US08_NEW_PUB_seq.*
5: /EMC_Celestra_SIDS3/ptodata/2/pubpna/PCT_NEW_PUB_seq.*
6: /EMC_Celestra_SIDS3/ptodata/2/pubpna/US10_NEW_PUB_seq.*
7: /EMC_Celestra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB_seq.*
8: /EMC_Celestra_SIDS3/ptodata/2/pubpna/US60_NEW_PUB_seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	118	5.1	1482	6	US-10-449-902-14217 Sequence 14217, A
C 2	114.5	5.0	1470	6	US-10-449-902-21879 Sequence 21879, A
C 3	114.5	5.0	2822	6	US-10-449-902-7723 Sequence 7723, Ap
C 4	113.5	4.9	2711	7	US-11-145-307A-189 Sequence 189, App
C 5	109.5	4.8	5193	7	US-11-270-040-5 Sequence 5, Appl
C 6	101.5	4.4	1945	7	US-11-293-697-1957 Sequence 1957, Ap
C 7	101	4.4	2259	7	US-11-217-529-274 Sequence 274, App
C 8	99	4.3	3287	6	US-10-449-902-12801 Sequence 12801, A
C 9	98.5	4.3	14756	6	US-10-505-928-570 Sequence 570, App

10	97	4.2	1502	7	US-11-297-134-55	Sequence 55, Appl
11	97	4.2	1619	6	US-10-953-349-34901	Sequence 34901, A
12	96	4.2	2010	6	US-10-449-902-20112	Sequence 20112, A
13	95.5	4.1	2293	6	US-10-449-902-26929	Sequence 26929, A
14	95	4.1	1488	6	US-10-449-902-24588	Sequence 24588, A
15	95	4.1	3731	6	US-10-449-902-27838	Sequence 27838, A
16	94.5	4.1	2329	7	US-11-293-697-1489	Sequence 1489, Ap
17	94.5	4.1	3066	6	US-10-449-902-12659	Sequence 12659, A
18	94.5	4.1	3773	7	US-11-293-697-841	Sequence 841, App
19	94.5	4.1	4239	6	US-10-511-937-543	Sequence 543, App
C 20	94.5	4.1	12225	6	US-10-501-834-1	Sequence 1, Appl
C 21	94	4.1	1221	7	US-11-284-978-14	Sequence 14, Appl
C 22	93	4.0	1948	1	US-09-949-925-71	Sequence 71, Appl
C 23	92.5	4.0	6152	6	US-10-528-658-6	Sequence 6, Appl
C 24	91.5	4.0	1720	6	US-10-449-902-20120	Sequence 20120, A
C 25	91.5	4.0	3317	7	US-11-293-697-1085	Sequence 1085, Ap
C 26	91.5	4.0	3646	6	US-10-511-937-358	Sequence 358, App
C 27	91	4.0	2127	6	US-10-953-349-8294	Sequence 8294, App
C 28	91	4.0	2168	6	US-10-449-902-16137	Sequence 16137, A
C 29	91	4.0	2295	7	US-11-293-697-946	Sequence 946, App
C 30	91	4.0	2970	7	US-11-293-697-1467	Sequence 1467, App
C 31	90.5	3.9	1339	6	US-10-449-902-1494	Sequence 1494, Ap
C 32	90.5	3.9	1450	6	US-10-449-902-22017	Sequence 22017, A
C 33	90.5	3.9	1526	6	US-10-449-902-13489	Sequence 13489, A
C 34	90	3.9	1719	6	US-10-449-902-26664	Sequence 26664, A
C 35	90	3.9	1987	7	US-11-293-697-1200	Sequence 1200, Ap
C 36	90	3.9	2408	7	US-11-293-697-315	Sequence 315, App
C 37	90	3.9	3069	6	US-10-449-902-12653	Sequence 12653, A
C 38	90	3.9	13128	7	US-11-118-524-1	Sequence 1, Appl
C 39	90	3.9	138941	6	US-10-489-730-10	GENERAL INFORMATION
40	89.5	3.9	1143	7	US-11-217-529-2750	Sequence 2750, Ap
41	89.5	3.9	1927	6	US-10-449-902-11403	Sequence 11403, A
42	89.5	3.9	2703	7	US-11-293-697-518	Sequence 518, App
43	89.5	3.9	2751	7	US-11-293-697-815	Sequence 815, App
44	89	3.9	2223	6	US-10-449-902-22201	Sequence 22201, A
45	89	3.9	2413	6	US-10-449-902-11507	Sequence 11507, A

ALIGNMENTS

RESULT 1
US-10-449-902-14217/c
; Sequence 14217, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A020511-US
; CURRENT APPLICATION NUMBER: US/10/449, 902
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14217
; LENGTH: 1482
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK064835
; DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-14217
Alignment Scores:
Pred. No.: 0.0336 Length: 1482
Score: 118.00 Matches: 74
Percent Similarity: 34.9% Conservative: 30
Best Local Similarity: 24.8% Mismatches: 116

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 14, 2006, 14:19:34 ; Search time 1438.58 Seconds
(without alignments)
5637.400 Million cell updates/sec

Title: US-10-601-319-2

Perfect score: 2302

Sequence: 1 MKAIIPLSLILPLTPQSA.....IVNEARIPACSLRSHHHHH 440

Scoring table:

	BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0			0.5
Fgapop 6.0			7.0
Delop 6.0			7.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

```

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abs/ABSSWEB_spool/US10601319/runat_12062006_180103_27735/app_query.fasta_1
-DB=Published Applications_NA_Main -QMT=fastcap -SUFFIX=rnpbm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs06p
-USER=US10601319 @CGN_1_1_3524 @runat_12062006_180103_27735 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORE=0 -WAIT -DSPBLOCK=100 -LONGLOG -DBV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

Database : Published Applications NA Main:

ID	Query	Length	DB	ID	Description
1	2302	100.0	1323	3	US-09-777-566A-1

ALIGNMENTS

RESULT 1

US-09-777-566A-1

Sequence 1, Application US/09777566A

Patent No. US20010055788A1

GENERAL INFORMATION:

APPLICANT: DIVERSA CORPORATION

APPLICANT: KORTZ, Jay

APPLICANT: KORTZ, Keith

TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF

FILE REFERENCE: DIVER1370-6

CURRENT APPLICATION NUMBER: US/09/777,566A

CURRENT FILING DATE: 2001-06-11

PRIOR APPLICATION NUMBER: US 09/318,528

PRIOR FILING DATE: 1999-05-25

PRIOR APPLICATION NUMBER: US 09/291,931

PRIOR FILING DATE: 1999-04-13

PRIOR APPLICATION NUMBER: US 09/259,214

PRIOR FILING DATE: 1999-03-01

PRIOR APPLICATION NUMBER: US 08/910,798

PRIOR FILING DATE: 1997-08-13

NUMBER OF SEQ ID NOS: 4

SOFTWARE: Patentin version 3.0

SEQ ID NO 1

LENGTH: 1323

TYPE: DNA

```
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1320)
; NAME/KEY: misc_feature
; LOCATION: (1)..(1323)
; OTHER INFORMATION: n is any nucleotide
US-09-777-566A-1

Alignment Scores:
Pred. No.: 1,32e-261 Length: 1323
Score: 2302.00 Matches: 440
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-601-319-2 (1-440) x US-09-777-566A-1 (1-1323)

Qy 1 MetLysAlaIleLeuLleProPheLeuSerLeuLleProLeuThrProGlnSerAla 20
Db 1 ATGAAGAGCATCTTAATCCCATTTTATCTCTTCTGATTCGGTTAAACCCCGCAATCTGCA 60
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db 61 TTCGCTCAGAGTGAGCGGAGCTGAGCTGGAAGTGTGGTGTGTCAGTCTGTCATGGT 120
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 121 GTGGCTGCTCAACCAAGGCCACGCAACTGATGAGGATGTACCCCGCAGACGCGATGCCA 180
Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuLleAlaTyrLeu 80
Db 181 ACCTGGCCGGTAAACATGGGTGGCTGACACCGCGGGTGGTGAAGTAACTGCGCTATCTC 240
Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLleLysLysGlyCysPro 100
Db 241 GGACATTAACCAACCCAGCGCTGTGTAGCCGAGATGTGGCGAAAGAGGCTGCGCG 300
Qy 101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 301 CAGTCTGGTCAGGTCGCGATATTATGCTGATGTCGACGAGCGTACCCGTAACACAGGCGAA 360
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCCTTCGCCCGCGGCTGGCACCTGACTGTGCAATACCGTACATACCCAGCGAGATACG 420
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCCAGTCCCGATCCGTTATTTATCTCTAATAACTGGCGTTGCCAACTGGATTAACGCG 480
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGACTGACGGATCCTCAGCAGGCGAGGAGGTCATTTGCTGACTTTTACCGGCGAT 540
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 541 CGGCAAAACGGCGTTTCGCGAATCGGAAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAAACGTGAGAAACAGGAGCAAGCTGTTCATTAACGCGAGGATTAACCATCGGAATC 660
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGAGCGCGCATATGCTCTAATTAACCGGTGCGGTAGCCCTCGCATCAATGCTGACG 720
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Db 721 GAGATATTTCTCTCGCAACAGCAAGCAAGGATGTCGCGAGCGGGGTGGGAAGGATCACC 780
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 841 CGCACCCAGAGGTTCGCCGACCGCGCCACCCCGTTATTGGATTGATCATGGCAGCG 900
Qy 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 901 TTGACGCCCATCCACCGCAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTACTG 960
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 961 TTTATTGCGGACACGATATACTTGGCAAAATCTCGCGCGCGCACTGGAGCTCAACTGG 1020
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlyLeuValPheGluArgTrp 360
Db 1021 ACCTTTCCCGGTGAGCGGATAACACCGCCAGGTTGTAAGTGGTGTGTAACCGCTGG 1080
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1081 CGTCGGCTAAGCATTAACAGCCAGTGGATTGAGTTTCGTTGCTTCCAGACTTTACAG 1140
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 1141 CAGATCGTGTATAAACCGCGCTGTCTATTAAATACGCGCGCGAGAGGTGAAACTGACC 1200
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1201 CTGCGCAGGATGTGAAGACGCAATGCGCAGGCGATGTTGCTGTGGCAGGTTTACGCA 1260
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
Db 1261 ATCGTGAATGACACGATACCGCGGTGCGATTTGAGATCTCATCATCATCATCATCAC 1320

RESULT 2
US-09-866-379-1
; Sequence 1, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1323)
; OTHER INFORMATION: n is any nucleotide
; NAME/KEY: CDS
; LOCATION: (1)..(1323)
; OTHER INFORMATION:
```

US-09-866-379-1

Alignment Scores:

Pred. No.: 1.32e-261 Length: 1323
Score: 2302.00 Matches: 440
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-601-319-2 (1-440) x US-09-866-379-1 (1-1323)

Qy 1 MetLysAlaIleLeuLeuProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
Db 1 ATGAAGCGATCTTAATCCATTTTATCTCTCTGATTCGGTTAAACCCCGCAATCTGCA 60
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db 61 TTGGCTCAGATGAGCCGAGCTGAGCTGGAAAGTGGTGAATGTGTCAGTCGTCATGGT 120
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTriPro 60
Db 121 GTGGCTCTCCACCAAGCCACCACTGATGAGGATGTCAACCCAGACGATGGCCA 180
Qy 61 ThrTrpProValLysLeuGlyTriLeuThrProArgGlyGlyGluLeuIleAlaTyriLeu 80
Db 181 ACCTGGCCGGTAAACCTGGTGGCTGACACCGCGNGGTGGTGAATGCTAATGCGCTATCTC 240
Qy 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
Db 241 GGACATTAACCAACGCCAGCGTCTGTAGCCAGCGATGTGCTGGCGAAAGGGCTGGCCG 300
Qy 101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 301 CAGTCTGGTCAAGTCGCGATTTATGCTGATGTCCAGAGCGTACCGGTAAACAGGCGAA 360
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCCTTCGCGCGCGCTGGCACCTGACTGTGCAATACCGTACATACCAGGACAGATAGC 420
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCCAGTCCCGATCCGTTATTTAATCCCTCTAAACACTGGCGTTTCCCACTGGATAACGCG 480
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGACTGACGGATCTCTCAGCAGGAGAGGGTCAATGTCTGACTTTACCGGCAT 540
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 541 CGGCAACCGCGTTTCGCGAATCGAAGCGGTGCTTAATTTTCGCAATCAAACTTTGTC 600
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAAACGTGAGAAACAGGACGAAGCTGTTCATTAAACGCGAGGATTAACCATCGGAATC 660
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGAGCGCGCAATGTCTATTAAACCGTGGTAAAGCTTACCGATCAATGCTGACG 720
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTriPglyArgIleThr 260
Db 721 GAGATATTTCTCTGCAACAGCACAGGAATGCGGAGCGGGGTGGGAAGGATCACCC 780
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
Db 781 GATTACACACCACTGGATCTTCAAGTTTGCATAACCGCAATTTTATTTCTCTCAAA 840
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 841 CGCAGCCAGAGGTTGCCGAGCGCGCCACCCCGCTTATTTGATTTGATCATGGCAGCG 900
Qy 301 LeuThrProHisProGlnLysGlnAlaTyriGlyValThrLeuProThrSerValLeu 320

Db 901 TTGACGCCCATCCACCGCAAAAAACAGCGGTATGGTGTGACATTACCCACTTCAGTACTG 960
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 961 TTTATTCGGGACACGATACCTAACTCTGGCAATCTCGGCGGCGACCTGGAGCTCAACTGG 1020
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 1021 ACGCTTCCCGGTCAAGCCGATAACACGCCCGCAGTGGTGAACGTGGTGTGAACGCTGG 1080
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1081 CGTCGGCTAAGCGATAACACGACGACGTCAGTGGATTTCAGGTTCGTGGTCTTCCAGACTTTACAG 1140
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyGluValLysLeuThr 400
Db 1141 CAGATGGGTGATAAACCGCGCTGTCAATTAATACGCCCGCGGAGAGGTGAACACTGACC 1200
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGCGATGTTCGTGGCAGGTTTACGCAA 1260
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
Db 1261 ATCGTGAATGAAGCAGCAGCATAACCGCGTGCAGTTTGAGATCTCATCACCACCATCAC 1320

RESULT 3

US-10-034-985-1
; Sequence 1, Application US/10034985
; Publication No. US20030049815A1
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/10/034,985
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US/09/580,515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; NAME/KEY: misc feature
; LOCATION: (1)...(1323)
; OTHER INFORMATION: n = A,T,C or G
US-10-034-985-1

Alignment Scores:

Pred. No.: 1.32e-261 Length: 1323
Score: 2302.00 Matches: 440
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-601-319-2 (1-440) x US-10-034-985-1 (1-1323)

Qy 1 MetLysAlaIleLeuLeuProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
Db 1 ATGAAGCGATCTTAATCCATTTTATCTCTCTGATTCGGTTAAACCCCGCAATCTGCA 60

```
QY 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db 61 TTGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAGTGTGGTGAATGTTCAGTCTGTCATGGT 120
QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrrPro 60
Db 121 GTGGTGTCTCCACCAAGGCCACCGCAACTGATGCAGGATGTACCCCCAGACGCGATGCCA 180
QY 61 ThrTrpProValLysLeuGlyTrrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db 181 ACCTGGCCGCTAAACCTGGGTGTGGCTCACACCGCGNGGTGGTGAGCTGAATCGCTATCTC 240
QY 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 241 GGACATTTACCAAGCCAGCGCTGGTAGCCGACGGATGTGTGGCAAAAAGGGCTGCCCG 300
QY 101 GlnSerGlyGlnValAlaIleLeuAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 301 CAGTCTGTGAGTTCGGATATTGCTGATGTGCAGCGGCTACCCGTAAACAGGCGAA 360
QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCCTTCGCCCGCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCAGGCAGATACG 420
QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAla 160
Db 421 TCCAGTCCCGATCGGTATTAAATCCTCTAAAACTGGCGTTTGGCCAACTGGATAACGCG 480
QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlyClySerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGACTGACGGATCTCTCAGCAGGGCAGGAGGTCAATTGCTGACTTTACCGGCAT 540
QY 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 541 CGGCAAAACGGCGTTTCGCAACTGGAACGGGTGCTTAATTTTCGCAATCAAACTTGTGC 600
QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAAACGTGAGAAACAGGACGAAGCTGTTCATTACCGCAGGCAATTACCATTGGAACTC 660
QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGAGCGCCGACATGCTCATTAACCGGTGCGGTAGCCTCGCATCAATGCTGACG 720
QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrrPglyArgIleThr 260
Db 721 GAGATATTCTCTCTGCAACAAGCACAGGGAATGCCGAGCGCGGTGGGAAGGATCACCC 780
QY 261 AspSerHisGlnTrrAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 781 GATTCACACAGTGGAACACCTTGCTAAGTTTGATTAACCGCAATTTTATTGTCTACAA 840
QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 841 CGCAGCGCCAGAGGTTCGCCGACGCGCGCCACCCCTTATGGATTGGATTCATGGCAGCG 900
QY 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 901 TTGACGCCCCATCCACGCAAAAACAGGCGGTATGGTGTGACATTAACCCCACTTCAGTACTG 960
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrr 340
Db 961 TTTATTCCCGGACACGATACTAATCTGGCAAACTCGCGCGCGCACCTGGAGCTCAACTGG 1020
QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrr 360
Db 1021 ACGCTTCGGGTGAGCCGGATAACACCGCCAGGTGTGTGAATGGTGTGTGAAGCTGG 1080
QY 361 ArgArgLeuSerAspAsnSerGlnTrrPileGlnValSerLeuValPheGlnThrLeuGln 380
Db 1081 CGTGGCTAAGCGGATAACAGCCAGTGGATTTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1140
QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
```

```
Db 1141 CAGATCGGTGATAAAACCGCGCTGTCTAATAATACCGCCCGGAGAGGTGAACCTGACC 1200
QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGCGCATGTGTTCGTTGGCAGGTTTTACGCAA 1260
QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
Db 1261 ATCGTGAATGAAGACGACGATACCGCGGTGCAGTTTGAGATCTCATCACCATCACCATCAC 1320

RESULT 4
US-10-430-356-1
; Sequence 1, Application US/10430356
; Publication No. US20030232041A1
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/10/430,356
; CURRENT FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: US/09/580,515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(1323)
; OTHER INFORMATION: n = A,T,C or G
US-10-430-356-1

Alignment Scores:
Pred. No.: 1,32e-261 Length: 1323
Score: 2302.00 Matches: 440
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 7 Gaps: 0

US-10-601-319-2 (1-440) x US-10-430-356-1 (1-1323)
QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAAGCGATCTTAATCCCATTTTTTATCTCTTCGATTCGTTAACCCCGCAATCTGCA 60
QY 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db 61 TTGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAGTGTGGTGAATGTTCAGTCTGTCATGGT 120
QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrrPro 60
Db 121 GTGGTGTCTCCACCAAGGCCACCGCAACTGATGCAGGATGTACCCCCAGACGCGATGCCA 180
QY 61 ThrTrpProValLysLeuGlyTrrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db 181 ACCTGGCCGCTAAACCTGGGTGTGGCTCACACCGCGNGGTGGTGAGCTGAATCGCTATCTC 240
QY 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
```

```
Db 241 GGACATTACCAACGCGCAGCGTCTGTAGCCAGCGATTGCTGGCGAAAGGGCTGCCCG 300
Qy 101 GlnSerGlyGlnValAlaIleAlaAspValLeuArgThrArgGlyThrGlyGlu 120
Db 301 CAGTCGTGTCAGTCGCGGATTTCTGATGTGCGACGAGCGTACCCGTAAACAGGCGAA 360
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCCTTCGCGCGCGGCTGGCACTGACTGTGCAATAACCGTATACACCAGCGAGATACG 420
Qy 141 SerSerProAspProLeuPheAsnProLeuYsThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCCAGTCGCGATCCGTTTAAATCCTCTAAACACTGGCGTGTGCAACTGGATAACGCG 480
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGATGACCGGATCCTCAGCAGGCGAGAGGTCAATTTGCTGCACTTTACCGGGCAT 540
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 541 CGGCAACCGCGGTTTCGGAACTGGAACGGGTGCTTAATTTCCGCATCAACTTGTGC 600
Qy 201 LeuYsArgGluYsGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGCGAGGCAATTACCATCGGAATC 660
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGACGCGGCAATGTCTCAATTAACCGGTGCGGTGAGCCTGCGCATCAATGCTGACG 720
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
Db 721 GAGATATTTCTTCTGCAACAGCACAGGAGATGCGGAGCGCGGGTGGGAGAGATCAAC 780
Qy 261 AspSerHisGlnTyrAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 781 GATTCAACACAGTGGAAACACCTTGCTAAGTTTGCATAACGCGCAATTTATTTGCTACA 840
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 841 CGCACGCGCAGAGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Qy 301 LeuThrProHisProProGlnYsGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 901 TTGACGCGCCCATCCACCGCAAAACAGCGGTATGTTGATGACATTACCCACTTCAGTACTG 960
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTyr 340
Db 961 TTTATTGCGGACACGATACTAATCTGGCAAAATCTCGCGCGCGCACCTGGAGCTCAACTGG 1020
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlyLeuValPheGluArgTyr 360
Db 1021 ACCTTCCTCGGTGAGCGGATAACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
Qy 361 ArgArgLeuSerAspAsnSerGlnTyrIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1081 GTCGCGCTAACGGATTAACGACGAGGTGATCAGGTTTCGCTGGTCTTCCAGACTTTACG 1140
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 1141 CAGATGCGTGATAAAACCGCGCTGTCTAATAATACCGCGCGCGAGAGGTGAACCTGACC 1200
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1201 CTGGCAGGATGTGAGAGCGAAATCGCGAGGCGCATGTGCTGTTGGCAGGTTTACGCAA 1260
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
Db 1261 ATCGGAATGAGCACCGCATACCGCGGTGAGTTTGATCTCATCACCACCATCATCAC 1320
```

RESULT 5

US-10-601-319-1

```
; Sequence 1, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT PHYASES AND METHODS OF MAKING
; TITLE OF INVENTION: AND USING THEM
; FILE REFERENCE: 09010-029011
; CURRENT APPLICATION NUMBER: US/10/601,319
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 216
; OTHER INFORMATION: n = A, T, C or G
US-10-601-319-1

Alignment Scores:
Pred. No.: 1,32e-261 Length: 1323
Score: 2302.00 Matches: 440
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: Gaps: 0

US-10-601-319-2 (1-440) x US-10-601-319-1 (1-1323)
Qy 1 MethLysAlaIleLeuIleProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
Db 1 ATGAAGCGGATCTTAATCCCATTTTATCTCTCTGATTCGTTAACCCGCAATCTGCA 60
Qy 21 PheAlaGlnSerGluProGluLeuYsLeuGluSerValValIleValSerArgHisGly 40
Db 61 TTCGCTCAGAGTAGCGCGGAGCTGAAGCTGAAAGTGTGGTGAATTGTCAGTCTCATGGT 120
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTyrPro 60
Db 121 GTGGTGCTTCACCAAGGCCACGCACTGATGTGAGGATGTCAACCCAGACGATGGCCA 180
Qy 61 ThrTyrProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db 181 ACCTGGCGGTAAACATGGGTGTGCTGACACCGCGNGGTGGTGGAGTAAATCGCTATCTC 240
Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
Db 241 GGACATTACCAACGCGCAGCGTCTGGTAGCCAGCGATTGCTGGCGAAAGGGCTGCCCG 300
Qy 101 GlnSerGlyGlnValAlaIleAlaAspValLeuArgThrArgGlyThrGlyGlu 120
```

```
Db 301 CAGTCTGGTCAGTCCGATATTATTCGTGATGTGACAGCGGTACCCGTAATAACAGCGCAA 360
Qy 121 AlaphealalaglyLeuAlaProAspCysAlaileThrValHisThrGlnAlaAspThr 140
Db 361 GCCTTCGCCCGCGGCTGGCACCTGACTGTGCAATACCGTACATACCCAGGCAGATACG 420
Qy 141 SerSerProAspProLeuPheAsnProLeuIysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCCAGTCCCGATCCGTTATTAATCCTCTAAAACTGGCGTTGGCCAACTGGATAACGGC 480
Qy 161 AsnValThrAspAlaileLeuSerArgAlaGlySerileAlaAspPheThrGlyHis 180
Db 481 AACGTGACTACGGCATCTCTACGAGGCGAGAGGGTCAATTGCTGATTTTACCGGCAT 540
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 541 CGGCAAAACGGGTTTCGCGAACTGGAACGGGTGCTTAATTTTCGCAATCAAACTTGTGC 600
Qy 201 LeuIysArgGluIysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAAACCGTGAGAAACAGGACGAAGCTGTTCAATTAACGAGGCATTAACCATCGGAATC 660
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyValAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGAGCGCGCAATGCTCTAATTAACGGTGGGTAAAGCTTCGCATCAATGCTGACG 720
Qy 241 GluilePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgileThr 260
Db 721 GAGATATTTCTCTGCAACAGCACAGGAAATGCGGAGCGCGGGTGGGGAAGATCAC 780
Qy 261 AspSerHisGlnTyrAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 781 GATTACACACGAGTGAACACCTTGCTAAGTTTGCATTAACGCGCAATTTTATTGCTACAA 840
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuileMetAlaAla 300
Db 841 CGCACGCGCAGAGGTGCGCGCAGCGCGCCACCCGTTATGGATTTGATCATGGCAGCG 900
Qy 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 901 TTGACGCGCCCATCCACCGCAAAACAGCGCGTATGGTGTGACATTAACCCACTTCAGTACTG 960
Qy 321 PheilealaglyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTyr 340
Db 961 TTTATTCGCGACACGATCTAATCTGCGAAATCTCGCGCGCGCAGCTGGAGCTCAACTGG 1020
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTyr 360
Db 1021 ACGTTTCCCGGTACCGCGATTAACACCGCGCAGGTGGTGNACTGGTGTGAAACGCTGG 1080
Qy 361 ArgArgLeuSerAspAsnSerGlnTyrileGlnValSerLeuValPheGlnThrLeuGln 380
Db 1081 CGTCGGCTAAGCGATAACAGCAGTGGATTTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1140
Qy 381 GlnMetArgAspIysThrProLeuSerLeuAsnThrProProGlyGluValIysLeuThr 400
Db 1141 CAGATGGGTGATAAACCCTGCTGCTAATTAATAACGCGCGCGAGAGGTGAACACTGACC 1200
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1201 CTGGCAGGATGTGAAGAGCGAATTCGACAGGCATGTTGCTGGCAGGTTTTACGCCAA 1260
Qy 421 IleValAsnGluAlaArgileProAlaCysSerLeuArgSerHisHisHisHisHis 440
Db 1261 ATCGTGAATGAAGCAGCATACCGCGTGCAGTTTGAGATCTCATCACCATCACCATCAC 1320
```

RESULT 6

US-10-933-115-1

; Sequence 1, Application US/10933115

; Publication No. US20050281792A1

; GENERAL INFORMATION:

; APPLICANT: Short, Jay M.

```
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 564462001822
; CURRENT APPLICATION NUMBER: US/10/933,115
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US/09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US/09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US/09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US/09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US/08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 216
; OTHER INFORMATION: n = A,T,C or G
US-10-933-115-1
```

Alignment Scores:

Pred. No.:	1,32e-261	Length:	1323
Score:	2302.00	Matches:	440
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	10	Gaps:	0

US-10-601-319-2 (1-440) x US-10-933-115-1 (1-1323)

Qy	1	MetLysAlaIleLeuileProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla	20
Db	1	ATGAAGCGATCTTAATCCATTTTATCTCTCTGATTCCGTTAACCCCGCAATCTGCA	60
Qy	21	PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly	40
Db	61	TTGCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGCTGATTGTCAGTCGTCATGTT	120
Qy	41	ValArgAlaProThrIysAlaThrGlnLeuMetGlnAspValThrProAspAlaTyrPro	60
Db	121	GTGCGTGTCTCAACCAAGGCCACGCAACTGATGAGGATGTCAACCCAGACGATGGCCA	180
Qy	61	ThrTyrProValIysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu	80
Db	181	ACCTGCCGCGTAACCTGGTTGGCTGTGACACCCGNGGTGGTGAGTAAATCGCTATCTC	240
Qy	81	GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaIysLysGlyCysPro	100
Db	241	GGACATTAACCAACGCGCAGCGTCTGTGTAGCCGACGGAATGCTGCGGAAAAAGGCTGCC	300
Qy	101	GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgIysThrGlyGlu	120
Db	301	CAGTCTGGTTCAGGTCCGGAATTTATGCTGATGTGACGAGCGTACCCGTAATAACAGCGCAA	360
Qy	121	AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr	140

```
Db 361 GCTTTTCGCGCGCGGCTGCGCCTCACTGTGCAATAACCGTATACATACCAGGACAGATACG 420
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCCAGTCCGATCCGTTATTAATCTCTAATAAATCGGTTGCAACTGGATACCGG 480
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGACTGACGCGATCTCAGCAGGCGAGGAGGTCAATTCGTGACTTTTACCGGCAT 540
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 541 CGCAAAACGCGCTTTCGGAACCTGAACGGGCTTAATTTTCCCAATCAAACTTTGTC 600
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTAAACCTGAGAAACAGACGAAAGCTGTTCAATTAACGCGCATTTACCATCGAATC 660
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGAGCGCGCACATGTCTCAATTAACCGTGGGTAAAGCTCGCATCAATGCTGACG 720
Qy 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
Db 721 GAGATATTTCTCTCTCAACAACAGCACAGGGAATGCCGAGCGCGGTGGGGAAGGATCAC 780
Qy 261 AspSerHisGlnThrAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 781 GATTACACACCTAGTGGACACCTTGTAGTTGTCATTAACGCGCAATTTATTTGCTACA 840
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 841 CGCAGCGCAGAGGTTGCCGCGAGCGCGCCACCGCTTATTTGGATTTTCATCATGCGCG 900
Qy 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 901 TTGACGCGCCCATCCACCCCAAAAACAGCGCGTATGGTGTGACATTAACCCACTTCAGTAC 960
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeuGluLeuAsnTrp 340
Db 961 TTTATTTGCGGACAGCATACTAATCTGCAATATCTCGCGCGGCACTGGAGCTCAACTGG 1020
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 1021 ACGTTTCCCGTTCAGCGGATAACACGCGCGCCAGGTGGTGAATGGTGTGTAACGCTGG 1080
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1081 CGTCGGCTTAACGCGATAACAGCGAGTGGATTTCAGGTTTCGCTTCCAGACTTTACAG 1140
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 1141 CAGATGGGTGATTAACCCCGCTGTCATTAATACGCCCGCCCGGAGAGGTGAACCTGACC 1200
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1201 CTGGCAGGATGTGAAGAGCGAAATCGCGAGGCGATGTGTTCTGTTGGCAGGTTTTACGCA 1260
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
Db 1261 ATCGTGAATGAACAGCATACCGCGCTGCAGTTTGAGATCTCATCACCATCACCATCAC 1320
```

RESULT 7

```
US-10-282-122A-7167
; Sequence 7167, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
```

```
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7167
; LENGTH: 1299
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-282-122A-7167
```

```
Alignment Scores:
Pred. No.: 1,12e-253 Length: 1299
Score: 2235.00 Matches: 430
Percent Similarity: 99.5% Conservative: 0
Best Local Similarity: 99.5% Mismatches: 2
Query Match: 97.1% Indels: 0
DB: 8 Gaps: 0
```

US-10-601-319-2 (1-440) x US-10-282-122A-7167 (1-1299)

```
Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAAGCGATCTTAATCCCAATTTTATCTCTCTGATTTCGTTAACCCCAATCTGCA 60
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db 61 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGGTGAATTGTTCAGTCGTCATGGT 120
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 121 GTGCGTCTCAACCAAGGCCACGCAACTGATGTCAGGATGTCAACCCAGACGATGGCA 180
Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db 181 ACCTGGCCGGTAAACCTGGGTGGCTGACACCGCGCGGTGTGAGCTAAATCGCCTATCTC 240
Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 241 GGACATTACCAACGCCAGCGCTCTGGTAGCCGCGATTGCTGGCGAAAAGGGCTGCCCG 300
Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 301 CAGTCTGGTTCGGTTCGGATTTATTTGCTGATGTGACGAGCGGTACCCGTAACAGGCGAA 360
```

```
QY 121 AlapheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCCTTCGCCCGCGCTGGACCTGACTGTGCATAACCGTACATCCAGGAGATACG 420
QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCCAGTCCCGCATCGTTATTATCTCTAAACCTGCGTTTGGCCAACCTGGATAACGCG 480
QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGACTACGCGATCCTCAGCAGGCGAGAGGGTCAATTGTCTGATTTTACCGGCAT 540
QY 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 541 CGGCAACCGGGTTTCGGCACTGGAAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600
QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAAACGTGGAACACAGGACGAAGCTGTTCTTAACGCGAGCATTTACCATCGGAATC 660
QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGAGCGCGACATGCTCATTAACCGGTGCGTAAGCTCGCATCAATGCTGACG 720
QY 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
Db 721 GAGATATTCTCTCGCAACAGCACAGGGAATGCCGAGCGGGGTGGGAAGATCACC 780
QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 781 GATTCACACAGTGGACACCTTGCTAAGTTTGATTAACGGGCNAATTTTATTTGCTACAA 840
QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 841 CGCACCGCAGAGGTGGCCGCGAGCGGCCACCCCGTTATTAGATTTGATCAAGACAGCG 900
QY 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 901 TTGACGCCCCATCCACCGCAAAAACAGCGGTATGGTGTGACATTTACCCACTTCAGTGCTG 960
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 961 TTTATCCCGGACACGATACTAATCTGCGAAATCTCGCGCGGCGACTGGAGCTCAACTGG 1020
QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlnLeuValPheGluArgTrp 360
Db 1021 ACGTCTCCCGTCAAGCGGATAAACACGCGCGAGGTGGTGAACCTGGTGTGTTGAACGCTGG 1080
QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1081 CGTCGGCTAAGCGATAACAGCAGTGGATTCAGGTTTCGGTGTCTTCCAGACTTTTACAG 1140
QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 1141 CAGATGCGTGATAAAACCGCGCTGCTCATTAATAACGCGCGCGGAGAGGTGAAACTGACC 1200
QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTCTGTTGGCAGGTGTTTACGCA 1260
QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1261 ATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTG 1296
```

RESULT 8

US-09-866-379-7

; Sequence 7, Application US/09866379

; Patent No. US20020136754A1

; GENERAL INFORMATION:

; APPLICANT: DIVERSA CORPORATION

; APPLICANT: SHORT, Jay

; APPLICANT: KRETZ, Keith

```
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
```

```
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-7
```

```
Alignment Scores:
Pred. No.: 1,95e-253 Length: 1901
Score: 2235.00 Matches: 430
Percent Similarity: 99.5% Conservative: 0
Best Local Similarity: 99.5% Mismatches: 2
Query Match: 97.1% Indels: 0
DB: 3 Gaps: 0
```

US-10-601-319-2 (1-440) x US-09-866-379-7 (1-1901)

```
QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 188 ATGAAAGCATCTTAATCCATTTTATCTCTCTGATTCCGTTAACCCCGCAATCTGCA 247
QY 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db 248 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAAGTGGTGGTATTGCTAGTCGTCACTG 307
QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 308 GTGCGTGTCTCAACCAAGGCCACGCAACTGTATGTCAGAGATGTCAACCCAGCATGGCCA 367
QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaIleLeu 80
Db 368 ACCTGCCCGGTAAACCTGGGTTGGCTGACACCGCGNGGTGGTGGCTAATCGCCTATCTC 427
QY 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 428 GGACATTACCAACGCCAGCGCTCTGGTAGCCGAGGATTCCTGGCGAAAAGGGCTCGCCG 487
QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 488 CAGTCTGGTCAGGTCCGATATTGCTGATGTCAGCAGCGTACCCTACCGTAAACAGCGCAA 547
QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCCTTCGCCCGCGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 607
QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 608 TCCAGTCCCGATCCGTTATTATCTCTAAACAACTGGCGTTTGGCAACTGGATACGCG 667
QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
```

```

Db 668 AACGTGACTGAGCGCATCTCAGCAGGCGAGAGGTCATTTGCTGACTTTACCGGGAT 727
Qy 181 ArgGlnThrAlaPheArgGluLeuValArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 728 CGGCAACCGCGTTTCGGAACTGGAACGGGTGCTTAATTTTCCCAATCAAACTGTGC 787
Qy 201 LeuValArgGluLeuValArgGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 788 CTTAAACGTGAAACAGCAGCAAGCTGTTCAATTAACGCGAGCATTAACCATCGGAATC 847
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 848 AAGGTGAGCGCGCAATGTCATTAACCGGTGCGTAAGCCTCGCATCAATGCTGACG 907
Qy 241 GluLeuPheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyValGlyLeuThr 260
Db 908 GAGATATTTCTCTGCAACAGCAGCGGAATCGCGAGCGGGGTGGGAGAGATCAAC 967
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
Db 968 GATTACACACAGTGAACACCTTGTAAAGTTTGCAATACGCGCAATTTAATTTGCTACA 1027
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuLeuMetAlaAla 300
Db 1028 CGCACGCGCAGAGGTGCGCGCAGCGCGCCACCGCTTATTAGATTTGATCAAGACAGCG 1087
Qy 301 LeuThrProHisProProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
Db 1088 TTGAGCGCCCATCCACCGCAAAACAGCGGTATGGTGTGACATTAACCACTTCAGTGCTG 1147
Qy 321 PheLeuAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 1148 TTTATCGCGGACACGATACATACTGCAAAATCTCGCGCGCACCTGGAGCTCAACTGG 1207
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlyLeuValPheGluArgTrp 360
Db 1208 ACGCTTCCTCGGTACGCGGATTAACACGCGCGAGGTGGTGAACCTGGTGTGAAACGCTGG 1267
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1268 CGTCGGCTAAGGATACAGCGCAGTGGATTCAGGTTTCGCTGCTTCACAGCTTTACAG 1327
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValValLeuThr 400
Db 1328 CAGATGCGTGATAAAACCGCGCTGTCTAATAATACGCGCGCGAGAGGTGAACCTGACC 1387
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1388 CTGGCAGGATGTGAAGACGAAATGCGCAGGCGCATGTGTTGTTGGCAGGTTTTACGCAA 1447
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1448 ATCGTGAATGAAGCAGCATACCGCGGTGCAGTTTG 1483

```

RESULT 9

```

US-10-156-660-3
; Sequence 3, Application US/10156660
; Publication No. US20030103958A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson R.
; APPLICANT: Garrett, James B.
; APPLICANT: O'Donoghue, Eileen
; APPLICANT: Machur, Eric J.
; TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM
; FILE REFERENCE: 09010-029007
; CURRENT APPLICATION NUMBER: US/10/156,660
; PRIORITY FILING DATE: 2002-10-01
; PRIORITY APPLICATION NUMBER: US 09/866,379

```

```

; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (188)...(1483)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 403
; OTHER INFORMATION: n = A, T, C or G
US-10-156-660-3

```

Alignment Scores:

```

Pred. No.: 1,956-253 Length: 1901
Score: 2235.00 Matches: 430
Percent Similarity: 99.5% Conservative: 0
Best Local Similarity: 99.5% Mismatches: 2
Query Match: 97.1% Indels: 0
DB: 6 Gaps: 0

```

US-10-601-319-2 (1-440) x US-10-156-660-3 (1-1901)

```

Qy 1 MetLysAlaLeuLeuLeuProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
Db 188 ATGAAGCGATCTTAATCCCATTTTATCTCTCTGATTCGTTAAACCCCGCAATCTGCA 247
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40
Db 248 TTGCTCAGAGTAGCGCGAGCTGAAGCTGGAAGTAGTGGTGTGATTTGTCAGTCGTATGCT 307
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 308 GTGGCTGCTCAACCAAGGCCACGCAACTGATGAGGATGTCAACCCAGACGATGGCCA 367
Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
Db 368 ACCTGGCGGTAAACCTGGGTGGCTGACACCGCGNGGTGGTAGCTAATCGCCTATCTC 427
Qy 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 428 GGACATTACCACGCCAGCGCTCTGGTAGCCGAGATTTGTCGCGCAAAAGGGGTGCCCG 487
Qy 101 GlnSerGlyGlnValAlaIleLeuAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 488 CAGTCTGCTCAGGTGCGCATTTATTTCTGATGTGACGAGCGGTACCCGTAAACAGGCGAA 547
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCCTTGGCGCGCGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 607
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAla 160
Db 608 TCAGTCCCGCATCGTTTATTTAATTCCTTAAAAAATCTGGCGTTTGGCCAATCGAATACGCG 667
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlyCysSerIleAlaAspPheThrGlyHis 180
Db 668 AAGTGACTGACGGATTCCTCAGCAGGCGAGGGGTCAATTTGCTGATTTTACCGGCAT 727
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200

```

```
Db 728 CGGCAACGGGGTTTCGGAACTGGAAACGGGTCTTAATTTTCGCAATCAAACTTGTC 787
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 788 CTTAAACGCTGAGAAACAGGACGAAAGCTGTTTCACTTAACGAGGCAATTAACCATCGGAATC 847
Qy 221 LysValSerAlaAspAenValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 848 AGGTGAGCCCGCAATGCTCTAATAACCGGTGCGTAAGCTCGCATCAATGCTGACG 907
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Db 908 GAGATATTTCTCTGCAACAGCACAGGAAATCCGCGAGCGGGTGGGAGGATCAC 967
Qy 261 AspSerHisGlnTrpAenThrLeuLeuSerLeuHisAenAlaGlnPheThrLeuLeuGln 280
Db 968 GATTACACAGGTGGAACACCTTGCTAAGTTTGCAATAACGCAATTTTATTGCTACAA 1027
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 1028 CGCACGCCAGAGGTTCGCCGAGCGCGCCACCCCGCTATTAGATTGATCAAGACAGCG 1087
Qy 301 LeuThrProHisProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
Db 1088 TTGAGCGCCCATCCACCGCAAAACAGCGGTATGTTGATACATTAACCATTTCACTGCTG 1147
Qy 321 PheIleAlaGlyHisAspThrAenLeuAlaAenLeuGlyGlyAlaLeuGluLeuAenTrp 340
Db 1148 TTTATCGCGGACACGATCTAATCTGCAAAATCTCGCGCGCGCACTGGAGCTCAACTGG 1207
Qy 341 ThrLeuProGlyGlnProAenAenThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 1208 ACGCTTCCCGGTGAGCGGATAACACCGCCAGGTGTGAATCGGTGTTGAACGCTGG 1267
Qy 361 ArgArgLeuSerAspAenSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1268 CGTCGGCTAAGCAATACACCGCATGTGATTCAGTTTCGTGCTCTTCCAGACTTTACAG 1327
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAenThrProProGlyGluValLysLeuThr 400
Db 1328 CAGATGCGTGATAAAACCGCGCTGTCAATTAATAACGCGCGCCGAGAGGTGAAACTGACC 1387
Qy 401 LeuAlaGlyCysGluGluArgAenAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1388 CTGCGAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTGCTGGCAGGTTTACGCAA 1447
Qy 421 IleValAenGluAlaArgIleProAlaCysSerLeu 432
Db 1448 ATCGTGAATGAGCACGCATACCGGCGTGCGAGTTTG 1483
```

RESULT 10

```
US-10-601-319-7
; Sequence 7, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
; FILE OF INVENTION: AND USING THEM
; FILE REFERENCE: 09010-029011
; CURRENT APPLICATION NUMBER: US/10/601,319
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
```

```
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (188)...(1483)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 403
; OTHER INFORMATION: n = A,T,C or G
US-10-601-319-7
```

Alignment Scores:

```
Pred. No.: 1,95e-253 Length: 1901
Score: 2235.00 Matches: 430
Percent Similarity: 99.5% Conservative: 0
Best Local Similarity: 99.5% Mismatches: 2
Query Match: 97.1% Indels: 0
DB: 8 Gaps: 0
```

US-10-601-319-2 (1-440) x US-10-601-319-7 (1-1901)

```
Qy 1 MetLysAlaIleLeuLeuProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 188 ATGAAGCGATCTTAATCCCAATTTTATCTCTTCGATTCGTTAACCCGCAATCTGCA 247
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db 248 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAAGTGGTGAATGTCAGTCTGATGTT 307
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 308 GTGGTGTCTCCAAACCAAGGCCACGCAACTGATGTCAGAGATGTCAACCCAGCGCATGCCA 367
Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
Db 368 ACCTGCCCGTAAAACTGGGTTGGCTGCACACCGCGNGGTGGTGAGCTAATCGCTATCTC 427
Qy 81 GlyHisTyxGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 428 GGACATTACCAACGCCGCGCTCTGGTAGCCGCGGATTGCTGGCGAAAAGGCTGCCCG 487
Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 488 CAGTCTGCTCAGTCGCGATTTATGCTGATGTCGACGAGCGTACCGTAAACAGCGGAA 547
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCCTTCGCCCGCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCGAGATCG 607
Qy 141 SerSerProAspProLeuPheAenProLeuLysThrGlyValCysGlnLeuAspAenAla 160
Db 608 TCAGTCCCGATCCGTTATTTAATCTCTAATAACTGGCGTTTGGCAACTGGATAACGCG 667
Qy 161 AenValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db 668 AACGTGACTGACGCGATCTCAGAGGGCAGGAGGGTCAATTGCTGACTTTTACCGGCA 727
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAenPheProGlnSerAenLeuCys 200
Db 728 CGGCAAAACGCGGTTTTCGCAACTGGAAACGGGTGCTTAATTTTCCGCAATCAAACTTTG 787
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
```

```

Db      788 CTTAAACGTGAGAAACAGGAGAAAGCTGTTTCATTAAACGCGAGCTATTACCATCGGAATC 847
Qy      221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db      848 AAGGTGAGCGCGGACAAATGCTCTATTACCGGTGGGTAAAGCTCGCATCAATGCTGACG 907
Qy      241 GlullePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTTPGlyArgileThr 260
Db      908 GAGATATTCTCTCTGCAACAAGACAGGGAATGCCGAGCGCGGGTGGGGAAGATCACC 967
Qy      261 AspSerHisGlnTTPAsnThrLeuLeuSerLeuHisAlaGlnPheTyrLeuLeuGln 280
Db      968 GATTACACCAAGTGAACACCTTGTCTAAGTTTGCATAACGCGCAATTTATTGCTACAA 1027
Qy      281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuLeuMetAlaAla 300
Db      1028 CGCAGCGCAGAGTTGCCCGCAGCGCGCCACCCCGTTATTAGATTGATCAAGACAGCG 1087
Qy      301 LeuThrProHisProGlnGlnGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db      1088 TTGACGCCCATCCACCGCAAAAACAGCGGTATGTTGATACATTACCCACTTCAGTGTG 1147
Qy      321 PheileAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuLeuLeuAsnTrp 340
Db      1148 TTTATCGCGGACAGGATACATACTGGCAATCTCGCGCGCGCACTGGAGCTCAACTGG 1207
Qy      341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db      1208 ACGTCTCCCGGTAGCGGATAACACGCGCGCAGGTGGTGAACCTGGTGTGAAACGCTGG 1267
Qy      361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db      1268 CGTCGGCTAAGCGATAACAGGACAGTGGATTTCAGGTTTCGCTGGCTTCCACACTTTAC 1327
Qy      381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyGluValPheGluArgTrp 400
Db      1328 CAGATGCGGTATAAAGCGCGCTGTCTATTAAATACGCGCGCGGAGAGGTGAACCTGACC 1387
Qy      401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db      1388 CTGCGAGGATGTGAAGAGCGAATTCGCGAGGCGATGTTTCGTTGGCAGGTTTACGAA 1447
Qy      421 IleValAsnGluAlaArgileProAlaCysSerLeu 432
Db      1448 ATCGTGAATGAAGCAGCATACCGCGCTGCAGTTTG 1483

```

RESULT 11

```

US-10-933-115-7
; Sequence 7, Application US/10933115
; Publication No. US20050281792A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Machur, Eric J.
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: 564462001822
; CURRENT APPLICATION NUMBER: US/10/933,115
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214

```

```

; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (188)...(1483)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 403
; OTHER INFORMATION: n = A,T,C or G
US-10-933-115-7

```

```

Alignment Scores:
Pred. No.: 1,95e-253 Length: 1901
Score: 2235.00 Matches: 430
Percent Similarity: 99.5% Conservative: 0
Best Local Similarity: 99.5% Mismatches: 2
Query Match: 97.1% Indels: 0
DB: 10 Gaps: 0

```

US-10-601-319-2 (1-440) x US-10-933-115-7 (1-1901)

```

Qy      1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
Db      188 ATGAAACGCGATCTTAATCCCATTTTATCTCTTCGTGATTACCGCCCAATCTGCA 247
Qy      21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40
Db      248 TTGCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGTGATGTGTCAGTCGTCATGGT 307
Qy      41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db      308 GTGCGTCTCCAAACCAAGGCCACGCACTGATGTCAGGATGTCCACCCAGACGCGATGGCA 367
Qy      61 ThrTrpProValLysLeuGlyTTPLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db      368 ACCTGGCCGTTAAACATGGGTTGCTGCTACACCGCGNGTGGTGAAGCTTAATCGCTATCTC 427
Qy      81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db      428 GGACATTACCAACGCCAGCGCTCTGTTAGCCGAGATTTGTCGCAAAAAGGGCTGCCCG 487
Qy      101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db      488 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTGACGAGCGGTACCCGTAACACAGGCGAA 547
Qy      121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db      548 GCCTTCGCCCGCGGCTGGCACCCTGCTGTGCAATTAACCGTACATACCCAGGCAATACG 607
Qy      141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db      608 TCAGTCCCGATCCGTTTATTAATCTCTTAAAACTGGCGTTTTCGCAACTGGATTAACGCG 667
Qy      161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db      668 AACGTGACTGACGCGATCTCTCAGCAGGCGAGGAGGTCAATTGCTGACTTTACCGGCGAT 727
Qy      181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db      728 CGGCAAAACGCGGTTTCGCAACTGGAACGCGGTCTTAATTTTCGCAATCAAACTTGTGC 787
Qy      201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db      788 CTTAAACGTGAGAAACAGGACGAAAGCTGTTTCAATTACCGAGCATTTACCATCGGAATC 847
Qy      221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240

```

```
Db 848 AAGGTGAGCGCCGACATGCTCATTAACCGGTGCGGTAAAGCTCGCATCAATGCTGACG 907
Qy 241 GluilePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgileThr 260
Db 908 GAGATATTTCTCTGCAACAGCACAGGAAATGCGGAGCGGGTGGGAGGATCACC 967
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerIleuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 968 GATTACACACGAGTGGAACTGCTTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA 1027
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 1028 CGACGCGCAGAGGTGCGCGAGCGCGCCACCCGCTTATAGATTGATCAACAGACGCG 1087
Qy 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 1088 TTGAGGCCCCCATCCACCGCAAAACAGCGGTATGTTGATGACATTTACCCACTTCAGTGCTG 1147
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 1148 TTTATCGCCGACACGATACTAATCTGSCAAATCTCGCGCGCGCACTGGAGCTCAACTGG 1207
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 1208 ACGTTCCCGGTAGCCGGATAACACCGCCAGGTGTGAACGTGGTTTGAACGCTGG 1267
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1268 CGTCGGCTAAGCGGATAACAGCAGTGGATTTCAGGTTTCGTTCCAGACTTTTACAG 1327
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 1328 CAGATGGGTGATAAAGCCGCTGTCTAATAACGCGCGCGGAGAGGTGAACCTGACC 1387
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1388 CTGGCAGGATGTGAAGACGGAAATGCGCAGCGCATGTTTCGTGGCAGGTTTACGCAA 1447
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1448 ATCGTGAATGAAGCACCATACCGCGCTGCAGTTTG 1483

RESULT 12
US-11-056-354-3
; Sequence 3, Application US/11056354
; Publication No. US20050246780A1
; GENERAL INFORMATION:
; APPLICANT: SHORT, Jay M.
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin A.
; APPLICANT: BARTON, Nelson R.
; APPLICANT: GARRETT, James B.
; APPLICANT: O'DONOGHUE, Eileen
; APPLICANT: MATHUR, Eric J.
; TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM AND METHODS
; FILE OF INVENTION: FOR MAKING AND USING THEM
; FILE REFERENCE: 564462001803/D1370-9C1
; CURRENT APPLICATION NUMBER: US/11/056,354
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 10/156,660
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
```

```
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli appA phytase
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (403)..(403)
; OTHER INFORMATION: n is any nucleotide
US-11-056-354-3

Alignment Scores:
Pred. No.: 1,95e-253 Length: 1901
Score: 2235.00 Matches: 430
Percent Similarity: 99.5% Conservative: 0
Best Local Similarity: 99.5% Mismatches: 2
Query Match: 97.1% Indels: 0
DB: 14 Gaps: 0

US-10-601-319-2 (1-440) x US-11-056-354-3 (1-1901)

Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 188 ATGAAAGCGATCTTAATCCATTTTATCTCTCTGATTTCCGTTAAACCCCGCAATCTGCA 247
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db 248 TTCGCTCAGAGTGAGCGGAGCTGGAAGCTGGAAAGTGTGGTGAATTGTCTAGTCGTCTAGT 307
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 308 GTGCGTGTCTCAACCAAGGCCACGCAACTGATGTCAGGATGTCAACCCAGACGATGGCCA 367
Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db 368 ACCTGCCCGTAAACTGGGTGGCTGACACCGCGNGTGGTGAGCTAATCGCTATCTC 427
Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 428 GGACATTTACCAACGCCAGCGTCTGTGTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCG 487
Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 488 CAGTCTGGTTCAGGTGCGGATTAATTGCTGATGTGCAGGCGTACCCGTAAACAGCGCGAA 547
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCCTTCGCCCGCGGGCTGGCACCTGACTGTGCANTAAACCGTACATACCCAGCGCATACG 607
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 608 TCCAGTCCCGATCCGTTATTTAATCTCTAAATAACTGGGGTTTGGCAACTGGATAACGCG 667
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db 668 AACGTGACTGACGCGATCTCTCAGCAGGCGAGGAGGTCAATTGCTGACTTTTACCGGCAT 727
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 728 CGGCAAAACGGCGTTTCGGNACTGGAAACGGGTGCTTAATTTTCCGCNATCAAACTGTGC 787
Qy 201 LeuLysArgLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 788 CTTTAAACGTGAGAAAACAGGACGAAAGCTGTTTCAATTAACGAGCATATTACCATCGAACTC 847
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 848 AAGGTGAGCGCGCAATGCTCTAATAACCGGTGCGGTAAAGCTCGCATCAATCGCTGACG 907
Qy 241 GluilePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgileThr 260
```

```
Db 908 GAGATATTCTCTCGCAACAGACACAGGGAATGCCGAGACCGGGGTGGGAGGATCACC 967
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 968 GATTACACACAGTGGAAACCTTGCTAGTTGTCATTAACGGCGCAATTTATTTGCTACA 1027
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuLeuMetAlaAla 300
Db 1028 CGCAGCCAGAGGTTGCCGCGAGCCGCCACCCCGTTATTAGATTGATCAAGACAGCG 1087
Qy 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 1088 TTGACGCCCATCCACCCGCAAAACAGCGCGTATGGTGTGACATTACCCACCTTCAGTGTG 1147
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 1148 TTTATCGCGGCACAGACTACTAATCTGGCAATCTCGCGGGCGACTGGAGCTCACTGG 1207
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlyLeuValPheGluArgTrp 360
Db 1208 ACCTTCCCGGTTCAGCCGATAACACGCCGCCGAGGTGGTGAACCTGGTGTGAAACGCTGG 1267
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1268 CGTCGGCTAAGCGATAACAGCCAGTGGATTGAGGTTTCGTGGTCTTCCAGACTTTACAG 1327
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 1328 CAGATGCGTGATAAAACCGCGCTGTCTAATAATAGCCGCCGCGGAGGTGAACTGACC 1387
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1388 CTGCAGGATGTGAAGAGCGAAATCGCGAGGCGATGTTGTTGGTGGAGGTTTACGCAA 1447
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1448 ATCGTGAATGAAGCAGCATATCCGCGTGCAGTTTG 1483
```

RESULT 13

```
US-10-601-319-6
; Sequence 6, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
; TITLE OF INVENTION: AND USING THEM
; FILE REFERENCE: 09010-029011
; CURRENT APPLICATION NUMBER: US/10/601,319
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
```

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 403..feature
; OTHER INFORMATION: n = A,T,C or G
US-10-601-319-6
```

```
Alignment Scores:
Pred. No.: 1,31e-252 Length: 1901
Score: 2228.00 Matches: 429
Percent Similarity: 99.3% Conservative: 0
Best Local Similarity: 99.3% Mismatches: 3
Query Match: 98.8% Indels: 0
DB: 8 Gaps: 0
```

US-10-601-319-2 (1-440) x US-10-601-319-6 (1-1901)

```
Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
Db 188 ATGAAGCGATCTTTAATCCCATTTTATCTCTCTGATTCGGTTAACCCTCCCAATCTGCA 247
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db 248 TTCCTCTCAGAGTAGCCGAGCTGAGCTGGAAAGTGTGTGATTGTTCAGTCTGTCATGCT 307
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 308 GTGCGTGTCTCCAAACCAAGGCCACGCAACTGATGATGAGGATGTCAACCCACAGCGCATGCCA 367
Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db 368 ACCTGGCGGTAAACCTGGGTGGCTGACACCGCGNGGTGGTGAGCTAATCGCTATCTC 427
Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 428 GGACATTACTGGCCCGCAGCGTCTGTAGCCACCGATTGCTGGCGAAAGGGCTGCCCG 487
Qy 101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 488 CAGTCTGTCAGGTCTCGCGATTATTGCTGATGTCACGAGCGTACCCGTAAACAGGCGAA 547
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCCTTCGCGCGCGGTGGCACCTGACTGTGTGCAATAACCCGTACATACCCAGCGAGATACG 607
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 608 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAACTGGCGTTTCCCACTCGATTAACCGCG 667
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db 668 AACGTGACTCGCGCATCTCTCAGCAGGCGCAGGAGGGTCAATTGCTGACTTTACCGGGCAT 727
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 728 CGGCAAAACCGCGTTTTCGCGAACTCGAAACGGGTGCTTAATTTTCGCGCAATCAAACTTGTGC 787
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 788 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTACGCGAGGATATTACCATCGGAATC 847
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 848 AAGGTGAGCGCGCAATGCTCATTAACCGGTGCGGTAGCCCTCGCATCAATCTCGACG 907
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Db 908 GAGATATTTCTCTGCAACCAAGCACAGGAAATCCGAGCGCGGGTGGGGAAGGATCACC 967
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 968 GATTACACAGTGGAAACCTTGTGCTAAGTTTGCATAACGCGCAATTTTATTTCTCTCAA 1027
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuLeuMetAlaAla 300
```

```

Db 1028 CGCAGCCAGAGGTTGCCCGCAGCGCCGCCACCCGTTATAGATTGATCAACAGACGCG 1087
Qy 301 LeuThrProHisProProGlnLysGlnAlaTyArgValThrLeuProThrSerValLeu 320
Db 1088 TTGACGCCCATCATCAGCGCAAAACAGCGGTATGGTGTGACATTTACCCACTTCAGTGCTG 1147
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 1148 TTTATCCCGGACACGATACCTAAATCTGGCAAAATCTCGCGCGCGCACTGGAGCTCAACTGG 1207
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluAtrTrp 360
Db 1208 ACGTCTCCCGGTACGCGGATAACACCGCGCGCGAGTGGTGAACCTGGTTTGAACGCTGG 1267
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1268 CGTCGGCTAAGCGGATAACAGCAGTGGATTTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1327
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 1328 CAGATGGCTGATAAAACCGCGCTGTCAATTAATAACGCGCGCGGAGAGGTGAACCTGACC 1387
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1388 CTGGCAGGATGTGAAGAGCGAATGCGCAGGCGATGTTGTTGTCGTCAGGTTTACGCAA 1447
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1448 ATCGTGAATGAAGCAGCATACCGCGCTGCAGTTTG 1483

```

RESULT 14

```

US-09-866-379-9
; Sequence 9, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Bileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER3170-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR FILING DATE: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR FILING DATE: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-9

```

Alignment Scores:
 Pred. No.: 2,27e-252 Length: 1901
 Score: 2226.00 Matches: 429
 Percent Similarity: 99.3% Conservative: 0

```

Best Local Similarity: 99.3% Mismatches: 3
Query Match: 96.7% Indels: 0
DB: 3 Gaps: 0

US-10-601-319-2 (1-440) x US-09-866-379-9 (1-1901)

Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 188 ATGAAGCGATCTTAATCCCATTTTATCTCTTCTGATTCCGTTAAACCCCGCAATCTGCA 247
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db 248 TTCGCTCAGAGTCAGCGGAGCTGAAGCTGGAAGTGTGTGATTGTTCAGTCGTATGGT 307
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 308 GTGCGTGTCTCAACACGAGCCACGCAACTGATGACGATGTCAACCCAGACGATGCGCA 367
Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyLeu 80
Db 368 ACCTGGCCGTTAAACTGGGTTGGCTGACACCCCGNGGTGGTGAAGCTAATCGCTATCTC 427
Qy 81 GlyHisTyrgLnaArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 428 GGACATTTACCAACGCCAGCGCTCTGGTAGCCGCGATTCGTCGCAAAAAGGGCTGCCCG 487
Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 488 CAGTCTGGTCAGGTCGCGATTAATTGCTGATGTGACGAGCGTACCCTGTAACACGAGCGAA 547
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCCTTCGCCCGCGCGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGAGATACG 607
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 608 TCCAGTCCCGATCCGTTTATTAATCCTCTAANAACCTGGCGTTTGGCAACTGGATAACGCG 667
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db 668 AACGTGACTGACGCGATCTCTCAGCAGGCGAGGAGGTCAATTGCTGACTTTTACCGGCGAT 727
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 728 CGCAAAACGCGGTTTCGCAACTGGAAACGGGCTTAATTTTCGCAATCAAACTTGTGC 787
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 788 CTTAAACGTGAGAAACAGGACGAAAGCTGTTTCAATTAACGCGAGCATTTACCATCGGAATC 847
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 848 AAGGTGAGCGCGCACAAATGTCATTAACCGGTGCGGTAAAGCCTGCATCAATGCTGACG 907
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Db 908 GAGATATTTCTCTCGCAACAGCACAGGGAATCCGAGAGCCGGGTGGGGAAGATCACC 967
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyLeuLeuGln 280
Db 968 GATTCACACCCAGTGAACACCTTTGCTAAAGTTTCATAACGCGCAATTTTGTGCTACAA 1027
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 1028 CGCACGCCAGAGGTTCGCCGACGCCCGCCACCGCTTATTAGATTGTGATCAAGACAGCG 1087
Qy 301 LeuThrProHisProGlnLysGlnAlaTyArgValThrLeuProThrSerValLeu 320
Db 1088 TTGACGCCCATCATCAGCGCAAAACAGCGGTATGGTGTGACATTTACCCACTTCAGTGCTG 1147
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 1148 TTTATCGCGGACACGATACCTAATCTGGCAAAATCTCGCGCGCGCACTGGAGCTCAACTGG 1207

```

```
QY 341 ThrLeuProGlyGlnProAspAenThrProProGlyGlyGluLeuValPheGluArgTrp 360
DB 1208 AGCTTCCTCCGGTCAGCGGATTAACACGCGCCAGGTGGTGAAGTGGTTGAAACGCTGG 1267
QY 361 ArgArgLeuSerAspAenSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
DB 1268 CGTCGGCTAAGCGGATAACAGCCAGTGGATTACAGTTTCAGTTTCGCTTCCAGACTTTACAG 1327
QY 381 GlnMetArgAspLysThrProLeuSerLeuAenThrProProGlyGlyValLysLeuThr 400
DB 1328 CAGATCGGTGATAAACCCTGTCATTAAATACGCCCGCGGAGAGGTGAATGACC 1387
QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
DB 1388 CTGGCAGGATGTGAAGAGCGAATCGCAGGCGCATGTGTTGCTTGGCAGGTTTACGCCAA 1447
QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
DB 1448 ATCGTGAATGAACGACGCATACCGCGGTGCAGTTTG 1483
```

RESULT 15

```
US-10-266-041-9
; Sequence 9, Application US/10266041
; Publication No. US20030072844A1
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: ENZYMES WITH IMPROVED PHYTASE ACTIVITY
; FILE REFERENCE: 19603/2791
; CURRENT APPLICATION NUMBER: US/10/266,041
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: US/09/540,149
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 60/127,032
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1489
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-266-041-9
```

Alignment Scores:

```
Pred. No.: 6 22e-252 Length: 1489
Score: 2221.00 Matches: 428
Percent Similarity: 99.1% Conservative: 0
Best Local Similarity: 99.1% Mismatches: 4
Query Match: 96.5% Indels: 0
DB: 6 Gaps: 0
```

US-10-601-319-2 (1-440) x US-10-266-041-9 (1-1489)

```
QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
DB 182 ATGAAGCGATCTTAATCCCATTTTATCTCTTTTGGATTCCGTTAAACCCCGCAATCTGCA 241
QY 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40
DB 242 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAGTGTGGTGAATGTTCAGCCGTCATGGT 301
QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
DB 302 GTGGTGCCGCCAACACAGCCACGCACTGATGAGGAGTGCACCCAGAGCGATGGCCA 361
QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
DB 362 ACCTGGCGGTAATAACTGGTTGGCTGACACCCAGCCGCTGGTGGTGAATTCGCTATCTC 421
QY 81 GlyHisThrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
DB 422 GGACATTACCAACGCGCAGCGTCTGGTGGCCGACGATTCGTTGGCGGAAAGGCGTCCCG 481
```

Search completed: June 14, 2006, 15:10:07
Job time : 1468.58 secs

```
QY 101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
DB 482 CAGCTGGTCAGGTCCGCGATTATTGCTGATGTCGACGAGCGTACCGGTAAACAGCGCNA 541
QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
DB 542 GCCTTCGCGCGCGGCTGGCACCTGACTGTGTCAATAACCGTACATPACCAGCAGATACG 601
QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
DB 602 TCCAGTCCCGATCCGTTATTATTCCTCTAAAAACTGGCGTTTGCCTCACTGGATAACCGG 661
QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
DB 662 AACGTGACTGACGCGATCCTCAGCAGCGCAGGAGGTCAATTTGCTGACTTTTACCGGCGAT 721
QY 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
DB 722 CGGCAAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCCAATTAAACTTTGTC 781
QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
DB 782 CTTAACCGTGAGAAACAGGACGAAAGCTGTTTCATTAAACGAGGCATTTACCATCGGAATC 841
QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
DB 842 AAGGTGAGCGCGGACCAATGTTTCATTAAACCGGTGGGTAAAGCTTCGCATCAATGCTGAGC 901
QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
DB 902 GAAATATTCTCTGCAACAGCAACAGGAAATGCGGAGCGGGGTGGGAAGGATCACT 961
QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuGln 280
DB 962 GATTACACACAGTGGAAACACCTTGTAAAGTTTGCATTAACGCGCAATTTTATTACTACAA 1021
QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
DB 1022 CGCAGCGCAGAGGTTCGCGCGAGTCCGCCACCCCGTTATTGGATTGATCATGGCAGCG 1081
QY 301 LeuThrProHisProProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
DB 1082 TTGACGCGCCCATCCACCGCAAAAAACAGCGGTATGCTGTGACATTTACCCACTTCAGTGTG 1141
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp 340
DB 1142 TTTATTGCGGACACGATACCTAATCTGGCAATCTCGCGCGGCACCTGGAGCTCAACTGG 1201
QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
DB 1202 ACGCTTCCAGGTACGCGGATAACACGCGCCAGGTGGTGAACCTGGTGTGNAACGCTGG 1261
QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
DB 1262 CGTCGGCTAAGCGGATAACAGCCAGTGGATTTCAGGTTCGCTGGTCTTCCAGACTTTACAG 1321
QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
DB 1322 CAGATGGGTGATAAACCCTGCTATCAATAAATACGCCCGCGGAGAGGTGAATCTGACC 1381
QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
DB 1382 CTGGCAGGATGTGAAGAGCGGAATGCGCAGGCGCATGTGTTGTTGGCGGTTTTTACGCCNA 1441
QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
DB 1442 ATCGTGAATGAAGCGCGCATACCGCGGTGCAGTTTG 1477
```

THIS PAGE LEFT BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 14, 2006, 12:41:52 ; Search time 224.541 Seconds
(without alignments)
5499.798 Million cell updates/sec

Title: US-10-601-319-2

Perfect score: 2302

Sequence: 1 MKAILIFLSLLIPTQSA.....IVNEARIPACSLRHHHHH 440

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abs/ABSSWEB_spool/US10601319/runat 12062006 180055 27578/app query.fasta_1
-DB=Issued_Patents_NA -QMT=fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=bloum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs03p -USER=US10601319 @CGN 1.1 381 @runat 12062006 180055 27578
-NCPU=6 -ICPU=3 -NO_WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /EMC_Celerra_SIDS3/ptodata/2/ina/1/1 COMB.seq.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/ina/5/5 COMB.seq.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/ina/6A/6A COMB.seq.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/ina/6B/6B COMB.seq.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/ina/7/7 COMB.seq.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/ina/H/COMB.seq.*
- 7: /EMC_Celerra_SIDS3/ptodata/2/ina/PCTUS COMB.seq.*
- 8: /EMC_Celerra_SIDS3/ptodata/2/ina/PP COMB.seq.*
- 9: /EMC_Celerra_SIDS3/ptodata/2/ina/RE COMB.seq.*
- 10: /EMC_Celerra_SIDS3/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2302	100.0	1323	3	US-09-259-214-1
2	2302	100.0	1323	3	US-09-318-528-1
3	2302	100.0	1323	3	US-09-291-931-1
4	2302	100.0	1323	3	US-09-580-515-1
5	2302	100.0	1323	3	US-09-866-379-1
6	2235	97.1	1901	3	US-09-866-379-7
7	2226	96.7	1901	3	US-09-866-379-9
8	2221	96.5	1489	3	US-09-540-149A-9

Alignment Scores:
Pred. No.: 1.05e-233 Length: 1323
Score: 2302.00 Matches: 440
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0

ALIGNMENTS

RESULT 1

US-09-259-214-1
; Sequence 1, Application US/09259214A

; Patent No. 6110719

; GENERAL INFORMATION:

; APPLICANT: Kretz, Keith

; TITLE OF INVENTION: NOVEL PHYTASE

; FILE REFERENCE: DIVER1370-1

; CURRENT APPLICATION NUMBER: US/09/259,214A

; CURRENT FILING DATE: 1999-03-01

; EARLIER APPLICATION NUMBER: 08/910,798

; EARLIER FILING DATE: 1997-08-13

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 1323

; TYPE: DNA

; ORGANISM: Escherichia coli

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)...(1320)

; NAME/KEY: misc feature

; LOCATION: (1)...(1323)

; OTHER INFORMATION: n = A,T,C or G

; US-09-259-214-1

```
DB: 3 Gaps: 0
US-10-601-319-2 (1-440) x US-09-259-214-1 (1-1323)
Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAGCGATCTTAATCCCACTTTTATCTCTTCGATTCCGTTAAACCCCGCATCTGCA 60
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db 61 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGGTGAATGTTCAGTTCGTCTGCT 120
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTTPPro 60
Db 121 GTGCGTCTCCAAACCAAGCGCCAGCAACTGATGAGGATGTCAACCCAGAGCATGGCCA 180
Qy 61 ThrTrpProValLysLeuGlyTTPLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db 181 ACCTGGCGGTAAAACTGGGTGGCTGACACCGCGGGTGGTGAATTAATCGCCTATCTC 240
Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 241 GGACATTAACCAACCCAGCGCTGTGTAGCCGACGATGTGCGCAAAAAGGGCTGCCG 300
Qy 101 GlnSerGlyGlnValAlaIleLeuAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 301 CAGTCTGGTCAGGTTCGGATTAATGCTGATGTGACGAGCGTACCGGTAAACAGGCGAA 360
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCCTTCGCGCGGGTGGCCACTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCCAGTCCCGATCCGTTAATTAATCTCTAAAACTGGCGTGTGCCAACTGGATAACGCG 480
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGACTCAGCGATCTCTCAGCAGGCGAGGAGGTCATTTGCTGATTTTACCGGCAT 540
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 541 CGCAACACGGGTTCGCAACTGGAACGGGTGCTTAATTTTCGCAATCAACTGTGC 600
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGAGGCAATTAACCATCGAACTC 660
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyValAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGAGCGCCGCAATGCTCATTAACCGGTGCGGTAAAGCTTCGCATCAATGCTGACG 720
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTTPGlyArgIleThr 260
Db 721 GAGATATTTCTCTGCAACAGCACAGGGAATCGCGAGCGCGGGTGGGAGGATCAC 780
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 781 GATTCACACGAGTGAACCTTGCTAAGTTTGCAACGCGCAATTTTATTTGCTACAA 840
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 841 CGCACGCCAGAGGTGGCCCGAGCGCGCCACCCCGTTATGGATTGATTCATGGCAGCG 900
Qy 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 901 TTGACGCCCCCATCCACCGCAAAACAGCGCGTATGGTGTGACATTAACCCACTTCAGTACTG 960
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp 340
Db 961 TTTATTTGCCGACACGATACATACTGTCGCAATCTCGCGCGCGCACTGGAGCTCAACTGG 1020
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 1021 ACCTTCCCGGTGAGCGGATACACCGCCAGGTGGTGAACCTGGTGTGAACGCTGG 1080
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1081 CGTCGGCTAAGCGATACACGACGAGTGGATTTCAGTTTCGCTGCTTCACAGACTTACAG 1140
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 1141 CAGATCGTGAATAAAACGCGCTGTCTAATAATACGCGCCCGGAGAGGTGAAACTGACC 1200
Qy 401 LeuAlaGlyCysGluGluArgAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1201 CTGCGAGGATGTGAAGAGCGAAATGCGAGGCGATGTGTTCGTGGCAGGTTTTACGCAA 1260
Qy 421 IleValAsnGluAlaAlaArgIleProAlaCysSerLeuArgSerHisHisHisHis 440
Db 1261 ATCGTGAATGAACGACGACATACCGGCGTGCAGTTTGAGATCTCATCACCATCACCAC 1320
RESULT 2
US-09-318-528-1
; Sequence 1, Application US/09318528
; Patent No. 6183740
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/318,528
; EARLIER FILING DATE: 1999-05-25
; EARLIER APPLICATION NUMBER: 09/291,931
; EARLIER FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (1)..(1320)
; NAME/KEY: misc_feature
; LOCATION: (1)..(1323)
; OTHER INFORMATION: n = A,T,C or G
US-09-318-528-1
Alignment Scores:
Pred. No.: 1,05e-233 Length: 1323
Score: 2302.00 Matches: 440
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0
US-10-601-319-2 (1-440) x US-09-318-528-1 (1-1323)
Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAGCGATCTTAATCCCACTTTTATCTCTTCGATTCCGTTAAACCCCGCATCTGCA 60
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db 61 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGGTGAATGTTCAGTTCGTCTGCT 120
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTTPPro 60
Db 121 GTGCGTCTCCAAACCAAGCGCCAGCAACTGATGAGGATGTCAACCCAGAGCATGGCCA 180
Qy 61 ThrTrpProValLysLeuGlyTTPLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db 181 ACCTGGCGGTAAAACTGGGTGGCTGACACCGCGGGTGGTGAATTAATCGCCTATCTC 240
Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 241 GGACATTAACCAACCCAGCGCTGTGTAGCCGACGATGTGCGCAAAAAGGGCTGCCG 300
Qy 101 GlnSerGlyGlnValAlaIleLeuAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 301 CAGTCTGGTCAGGTTCGGATTAATGCTGATGTGACGAGCGTACCGGTAAACAGGCGAA 360
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCCTTCGCGCGGGTGGCCACTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 420
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCCAGTCCCGATCCGTTAATTAATCTCTAAAACTGGCGTGTGCCAACTGGATAACGCG 480
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGACTCAGCGATCTCTCAGCAGGCGAGGAGGTCATTTGCTGATTTTACCGGCAT 540
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 541 CGCAACACGGGTTCGCAACTGGAACGGGTGCTTAATTTTCGCAATCAACTGTGC 600
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGAGGCAATTAACCATCGAACTC 660
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyValAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGAGCGCCGCAATGCTCATTAACCGGTGCGGTAAAGCTTCGCATCAATGCTGACG 720
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTTPGlyArgIleThr 260
Db 721 GAGATATTTCTCTGCAACAGCACAGGGAATCGCGAGCGCGGGTGGGAGGATCAC 780
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 781 GATTCACACGAGTGAACCTTGCTAAGTTTGCAACGCGCAATTTTATTTGCTACAA 840
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 841 CGCACGCCAGAGGTGGCCCGAGCGCGCCACCCCGTTATGGATTGATTCATGGCAGCG 900
Qy 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 901 TTGACGCCCCCATCCACCGCAAAACAGCGCGTATGGTGTGACATTAACCCACTTCAGTACTG 960
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp 340
Db 961 TTTATTTGCCGACACGATACATACTGTCGCAATCTCGCGCGCGCACTGGAGCTCAACTGG 1020
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
```

```

Db 181 ACCTGGCGGTAAACTGGTGGCTGACACCGCGNGGTGGTGAAGCTAATCGCCTATCTC 240
Qy GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
Db GGACATTACCAACGCGCAGCGTCTGGTAGCGACGATTTGCTGGCGAAAGGGCTGCCGG 300
Qy GlnSerGlyGlnValAlaLysLeuAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db CAGCTGCTGCTAGGTCCGATATTATCTGATGTGACGAGCGGTACCGTAAACAGCGCAA 360
Qy AlaPheAlaAlaGlyLeuAlaProAspCysAlaLysThrValHisThrGlnAlaAspThr 140
Db GCCTTCGGCGCGGGCTGGCAGCTGCTGTCATACCGATACCGATACCGAGCAGATACG 420
Qy SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db TCCAGTCCCGATCCGTTATTTAATCTCTAAAACTGGCGTTTGCCTGCACTTACCGGGCAT 480
Qy AsnValThrAspAlaLysSerHisArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db AACGTGACTGCGGATCCTCAGCAGGCGCAGGAGGTCAATTTGCTGACTTTACCGGGCAT 540
Qy ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db CGCMAACCGGCTTTCGGAACCTGGAACCGGTGCTTAATTTTCGCAATCAAACTTTGTC 600
Qy LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db CTAAACGTGAGAAACAGACGAGAAAGCTGTTCAATTAACGCGAGCATTTACCATCGGAATC 660
Qy LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db AAGGTGAGCGCGCAATGTCTCAATACCGGTGCGGTAAAGCCTCGCATCAATGCTGACG 720
Qy GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
Db GAGATAATTTCTTCGCAACAGCAGGGAATGCCGAGCGCGGTGGGGAAGGATCAC 780
Qy AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db GATTACACACAGTGGACACCTTGTCTAAGTTTGATACGCGCAATTTATTTGCTACAA 840
Qy ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db CGCAGCGCAGAGTTGCCCGCAGCGCGCCACCGCTTATTTGGATTTGATCATGCGACG 900
Qy LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db TTGACGCCCCATCCACCGCAAAAACAGCGGTATGCTGTGACATTTACCCACTTCAGTACTG 960
Qy PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db TTTATTTGCGGACAGCATCTAATCTGGCAATCTCGCGCGCGCACTGGAGCTCACTGG 1020
Qy ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db ACCTTCCCGGTACGCGGATAACACGCGCGCAGGTGGTGAACCTGGTTTGAACGCTGG 1080
Qy ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db CGTGGCTTAACCGGATAACAGCAGGTGGATTTCAGTTTTCGCTGGTCTTCCAGACTTTACAG 1140
Qy GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db CAGATGCTGTATAAACCGCGCTGTCTAATTAATACGCGCGCGGAGAGGTGAACTGACC 1200
Qy LeuAlaGlyCysGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db CTGCGAGGATGTGAAGACGAAATCGCGAGGCGATGTTGTTGTTGGCAGGTTTTACGCAA 1260
Qy IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
Db ATCGTGAATGAAGCAGCATACCGCGGTGAGTTTGAGATCTCATCATCCATCACCATCAC 1320

```

```

RESULT 3
US-09-291-931-1
; Sequence 1, Application US/09291931A
; Patent No. 6190897
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/291,931A
; CURRENT FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; NAME/KEY: misc_feature
; LOCATION: (1)...(1323)
; OTHER INFORMATION: n = A,T,C or G
US-09-291-931-1

Alignment Scores:
Pred. No.: 1.05e-233 Length: 1323
Score: 2302.00 Matches: 440
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-601-319-2 (1-440) x US-09-291-931-1 (1-1323)
Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAAGCGATCTTAATCCCATTTTATCTCTCTGATTCGGTTAACCCCAATCTGCA 60
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40
Db 61 TTCCTCAGAGTGAGCCGAGCTGAAGCTGGAAAGTGTGGTATGTCAGTCGTCATGGT 120
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 121 GTGCGTCTCCAAACCAAGGCCACGCAACTGATGCAGGATGTACCCAGACGATGGCCA 180
Qy 61 ThrTrpProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db 181 ACCTGGCGGTAAACCTGGGTGGCTGACACCGCGNGGTGGTGAAGCTAATCGCCTATCTC 240
Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
Db 241 GGACATTACCAACGCGCAGCGTCTGGTAGCGACGATTTGCTGGCGAAAGGGCTGCCCG 300
Qy 101 GlnSerGlyGlnValAlaLysLeuAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 301 CAGTCTGCTGAGTTCGCGATTAATTTGCTGATGTGACGAGCGGTACCCGTAACAGCGCAA 360
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaLysThrValHisThrGlnAlaAspThr 140
Db 361 GCCTTCGGCGCGGGCTGGCAGCTGCTGTCATACCGATACCGATACCGAGCAGATACG 420
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCCAGTCCCGATCCGTTATTTAATCTCTAAAACTGGCGTTTGCCTGCACTTACCGGGCAT 480
Qy 161 AsnValThrAspAlaLysSerLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180

```

```
Db 481 AACGTGACTGACGGATCTCTCAGCAGGCGAGGAGGGTCAATTGCTGACTTTTACCGGCGCAT 540
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 541 CGGCAACCGCGGTTTCGCGAAGTGGAAAGCGGTCTTAATTTTCGCGCAATCAAACTTGTGC 600
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAAACGTGAGAACAGGACGAAGCTGTTCATTAAACGAGGATTTACCATCGGAACCTC 660
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGAGCGCGACATGCTCTAATTAACCGGTGCGGTAAAGCTCGCATCAATGCTGACG 720
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
Db 721 GAGATATTCTCTCTGCAACAGCACAGGAAATGCGGAGCGCGGTGGGAAGGATCACCC 780
Qy 261 AspSerHisGlnTyrAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 781 GATTCAACACAGTGGAAACCTTCTGTAAGTTTGCATTAACGCGCAATTTTATTGCTACAA 840
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 841 CGCAGCGCAGAGGTTGCCGCGAGCGCGCCACCCCGTTATTGGAATTTGATCATGGCAGCG 900
Qy 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 901 TTGACGCGCCCATCCACCGCAAAACAGCGCGTATGGTGTGACATTTACCACCTTCAGTACTG 960
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTyr 340
Db 961 TTTATTTCGCGACACGATACTAATCTGGCAAACTCGCGCGCGCACCTGGAGCTCAACTGG 1020
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTyr 360
Db 1021 ACGCTTCCCGGTGAGCGGATAACACCGCCAGGTGTTGACTGGTGTGTTTGAACGCTGG 1080
Qy 361 ArgArgLeuSerAspAsnSerGlnTyrIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1081 CGTCGGCTAAGCGGATAACAGCAGTCAGTGTGATTCAGGTTTTCGCTGGTCTTCCAGACTTTACAG 1140
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 1141 CAGATGGCTGATAAAACCGCGCTGTCAATTAATACGCGCCCGGAGAGGTGAACCTGACC 1200
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1201 CTGGCAGGATGTGAAGAGCGGAATGCGCAGGCGCATGTTCGTTGGCAGGTTTTTACGCAA 1260
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
Db 1261 ATCGTGAATGAAGCACGACATACCGCGGTGCGAGTTTGAGATCTCATCAACCATCACCATCAC 1320
```

RESULT 4

```
US-09-580-515-1
; Sequence 1, Application US/09580515
; Patent No. 6720014
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/580,515
; PRIOR FILING DATE: 1999-05-25
; PRIOR FILING DATE: 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR FILING DATE: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR FILING DATE: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR FILING DATE: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
```

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; NAME/KEY: misc_feature
; LOCATION: (1)...(1323)
; OTHER INFORMATION: n = A,T,C or G
US-09-580-515-1

Alignment Scores:

Pred. No.: 1,05e-233 Length: 1323
Score: 2302.00 Matches: 440
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-601-319-2 (1-440) x US-09-580-515-1 (1-1323)

```
Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAAGCATCTTAATCCATTTTATCTCTCTGATCCGTTAACCCCGCAATCTGCA 60
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db 61 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGGTATTGTCACTCGTCATGGT 120
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTyrPro 60
Db 121 GTGCGTGTCTCAACCAAGGCCACGCAACTGATCAGCATGTCAACCCAGACGATGGCCA 180
Qy 61 ThrTyrProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db 181 ACCTGCGCGGTAAACATGGGTTGGCTGTGACACCGCGNGTGGTGAGCTAATCGCTATCTC 240
Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 241 GGACATTAACCAACGCCAGCGTCTGGTAGCCGCGATTTCTGGCGAAAAAGGGCTGCCG 300
Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 301 CAGTCTGGTCAGTTCGCGATTATTGCTGATGTGACGAGCGTACCCGTAACAGGCGAA 360
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCCTTCCGCGCGCGGTGGCACCTGACTGTGCAATTAACCGTATACCCAGGCGATACG 420
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCCAGTCCCGCATCCGTTATTATTAATCTCTAAAAAAGTGGCGTTTTGCCAACTGGATAACGCG 480
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlyCysIleAlaSerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGACTGACGCGATCTCTCAGCAGGCGAGGAGGTCAATTGCTGACTTTTACCGGCGAT 540
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 541 CGGCAACCGCGGTTTCGCGAAGTGGAAACGGGTGCTTAATTTTCGCGCAATCAAACTTGTGC 600
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAAACGTGAGAAACAGGACGAAGCTGTTTCAATTAACCGGTCGAGTATACCATCGGAACCTC 660
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGAGCGCGCACATGCTCATTAACCGGTGCGGTGAAGCTCGCATCAATGCTGACG 720
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
```

```
Db 721 GAGATATTCTCTCTCAACAACAGCAGAGGAATGCCGGAGCCGGGGTGGGAGGATCACC 780
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 781 GATTTCACACCACTGGACACCTTGTAGTTTGATTAACGGCAATTTATTTGCTACAA 840
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuMetAlaAla 300
Db 841 CGCAGCCAGAGGTTGCCCGCAGCGCGCACCCGGTTATTGGATTTCATGTCAGCG 900
Qy 301 LeuThrProHisProGlnLysGlnAlaTyrGlnValThrLeuProThrSerValLeu 320
Db 901 TTGACGCCCATCCACCGCAAAACAGCGGTATGTGTGACATTACCCACTTCAGTCTG 960
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 961 TTTATTGCGGACACGATCTAATCTGGCAATCTCGGGGGCACTGGAGCTCAACTGG 1020
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 1021 ACGCTTCCCGGTACGCCGATAACACGCCCGCCAGGTGGTGAACCTGGTTTGAACGCTGG 1080
Qy 361 ArgArgIleuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1081 CGTCGGCTTAAGCGATAACAGCCAGTGGATTCAAGTTTCGCTGCTTCCAGACTTTACAG 1140
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 1141 CAGATCGCTGATTAACACCGCGCTGTCTAATTAACGCCCGCCGGAGAGGTGAACCTGACC 1200
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1201 CTGGCAGGATGTGAAGAGCGAAATCGCGCAGGCGATGTGTCTGGTCAGGTTTACGCAA 1260
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHis 440
Db 1261 ATCGTGAATGAAGCAGCAGCATACCGCGCGTCAGTTTGAGATCTCATCACCATCACCATCAC 1320
```

RESULT 5

```
US-09-866-379-1
; Sequence 1, Application US/09866379
; Patent No. 6855365
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1323)
```

```
; OTHER INFORMATION: n is any nucleotide
; NAME/KEY: CDS
; LOCATION: (1)..(1323)
; OTHER INFORMATION:
US-09-866-379-1
```

```
Alignment Scores:
Pred. No.: 1,05e-233 Length: 1323
Score: 2302.00 Matches: 440
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0
```

US-10-601-319-2 (1-440) x US-09-866-379-1 (1-1323)

```
Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuProLeuThrProGlnSerAla 20
Db 1 ATGAAGCGATCTTAATCCCATTTTATCTCTCTGATTCCGTTAAACCCCGCAATCTGCA 60
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db 61 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAAGTGGTGATTGTTCAGTCGTCTATGCT 120
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 121 GTGCGTGTCTCCAAACCAAGGCCACCACTGATGCAGGATGTCAACCCAGACGCGATGCCA 180
Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db 181 ACCTGGCGGTAAACCTGGGTGGCTGACACCGCGNGGTGGTGAAGCTAATCGCTATCTC 240
Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLeuGlyCysPro 100
Db 241 GGACATTTACCAACCCAGCGTCTGTAGCCAGCGATTGCTGGCGAAAGGGCTGCCCG 300
Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 301 CAGTCTGGTCAGTTCGCGATTATTATCTCTAAACTGGCGTTTGCCTCAACTCGATTAACCG 360
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCCTTCGCGCGCGGCTGGCACCTGACTGTGTGCAATAACCGTACATACCCAGCAGATACG 420
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCCAGTCCCGATCCGTTATTATTATCTCTAAACTGGCGTTTGCCTCAACTCGATTAACCG 480
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGACTGCGCGATCTCTCAGCAGGCGAGGAGGTCATTTGCTGACTTTACCGGGCAT 540
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 541 CGGCAAAACCGCGTTTTCGCGAACTCGAAACCGGTGCTTAATTTTCCGCAATCAAACTGTGC 600
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCGAGGATTAACCATCGGAATC 660
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGAGCGCGCAATGCTCATTAACCGGTGCGGTGCGGTGAGCTCGCATCAATCTGACG 720
Qy 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Db 721 GAGATATTCTCTCTGCAACAGCACAGGGAATGCCGAGCGCGGGTGGGGAAGGATCAC 780
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuGln 280
Db 781 GATTTCACACCACTGGAAACCTTGTCTAAGTTTGATTAACCGCGCAATTTTATTTCTCTCAA 840
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
```

```
Db      841 CGCAGCCGAGAGTTGCCGAGCGGCCACCCCGTATTGGATTGGATCATGGCAGCG 900
Qy      301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db      901 TTGAGCGCCCATCCACCGCAAAACAGCGGTATGGTGTGACATTACCCACTTCAGTACTG 960
Qy      321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db      961 TTTATTGCGGACACGATCTAATCTGCAAACTCTCGCGCGCGCACTGGAGCTCAACTGG 1020
Qy      341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db      1021 ACGTTTCCCGTCAGCCGGATAACACGCGCCAGCGTGGTGAACCTGGTGTGAAACGCTGG 1080
Qy      361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db      1081 CGTCGGCTAAGCGATAACAGCCAGCTGGATTTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1140
Qy      381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db      1141 CAGATGGGTGATAAAACCGCGCTGTCTATTAAATACGCGCGCGGAGAGGTGAACCTGACC 1200
Qy      401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db      1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGCGCATGTGTCTGTTGGCAGGTTTTACCAG 1260
Qy      421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
Db      1261 ATCGTGAATGAAGCACGCATACCGCGGTGCAGTTTGAGATCTCATCACCATCAC 1320
```

RESULT 6

```
US-09-866-379-7
; Sequence 7, Application US/09866379
; Patent No. 6853365
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR FILING DATE: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-7

Alignment Scores:
Pred. No.: 2,34e-226 Length: 1901
Score: 2235.00 Matches: 430
Percent Similarity: 99.5% Conservative: 0
```

```
Best Local Similarity: 99.5% Mismatches: 2
Query Match: 97.1% Indels: 0
DB: 3 Gaps: 0
US-10-601-319-2 (1-440) x US-09-866-379-7 (1-1901)
Qy      1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db      188 ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCTTAACCCCGCAATCTGCA 247
Qy      21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db      248 TTGGCTCAGAGTGAAGCGGAGCTGAAGCTGGAAGTGTGGTGTGATTTGTCAGTCTCATGGT 307
Qy      41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db      308 GTGCGTGTCTCAACCAAGCGCACCGCAACTGATGCAGGATGTCAACCCAGACGCGATGCCA 367
Qy      61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db      368 ACCTGGCCGGTAAACCTGGGTTGGCTGACACCCCGNGGTGGTGAGCTTAATCGCCTATCTC 427
Qy      81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db      428 GGACATTACCAACGCCAGCGCTCTGGTAGCCGCGATTGCTGGCGAAAGAGGCTGCCCG 487
Qy      101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db      488 CAGTGTGGTTCAGGTTCGCGATTATTGCTGTGATGTGCAGCGGTACCCCGTAAACAGGCGAA 547
Qy      121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db      548 GCCTTGGCCCGCGGCTGGCACCTGACTGTGCAATAACCGTACATATCCAGGCGAGATACG 607
Qy      141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db      608 TCCAGTCCCGATCCGTTATTATCTCTTAAACCTGGCGTTGCGCAACTGGATAACGCG 667
Qy      161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db      668 AACGTGACTGACGCGATCTCTCAGCAGGCGAGGAGGTCAATTGCTGATTTTACCGGCGAT 727
Qy      181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLysCys 200
Db      728 CGGCAACCGCGGTTTCGCGAACTGGAAACGGGTCTTAATTTTCGCAATCAAACTGTGCG 787
Qy      201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db      788 CTTAAACGTGAGAAACAGGACGAAAGCTGTTTCAATTAACGCGAGGCAATTACCATCGAACTC 847
Qy      221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db      848 AAGGTGAGCGCCGACATGCTCTCAATTAACCGGTGCGGTAAAGCCCTCGCATCAATGCTGACG 907
Qy      241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Db      908 GAGATATTTCTCTGCAACAGACAGGGAATGCCGAGCGCGGGGTGGGAGGATCACC 967
Qy      261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db      968 GATTCAACACGAGTGAACACCTTGTCTAAGTTTGTCAATTAACGCGCAATTTTATTTCTCAAA 1027
Qy      281 ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db      1028 CGCAGCGCAGAGGTGCGCGCAGCGCCACCCCGCTTATTAGATTGATCAACAGACGCG 1087
Qy      301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db      1088 TTGACGCCCATCCACCGCAAAACAGCGGTATGGTGTGACATTTACCCACTTCAGTCTGTG 1147
Qy      321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db      1148 TTTATCGCGGACACGATACTAATCTGGCAAACTCTCGCGCGCGCACTGGAGCTCAACTGG 1207
```

QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrrp 360
DB 1208 ACGCTTCCGGTCAGCCGGGATAACACGCCCGCCAGTGGTGAACCTGGTGG 1267
QY 361 ArgArgLeuSerAspAsnSerGlnTrrpIleGlnValSerLeuValPheGlnThrLeuGln 380
DB 1268 CGTCGGCTAAGCGATAACAGCCAGTGGATTTCAGGTTTCGGTCTTCCAGACTTTACAG 1327
QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
DB 1328 CAGATCGCTGATAAACCCTGCTGTCATTAATACGCCCGCGAGAGGTGAACCTGACC 1387
QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
DB 1388 CTGCGCAGGATGTGAAGAGCGAAATCGCAGGCGCATGTGTTGGTGGCAGGTTTAAAGCAA 1447
QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
DB 1448 ATCGTGAATGAAGCAGCATACCGCGGTGCAGTTTG 1483

RESULT 7

US-09-866-379-9
; Sequence 9, Application US/09866379
; Patent No. 6855365
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 9
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-9

Alignment Scores:
Pred. No.: 2,1e-225 Length: 1901
Score: 2226.00 Matches: 429
Percent Similarity: 99.3% Conservative: 0
Best Local Similarity: 99.3% Mismatches: 3
Query Match: 96.7% Indels: 0
DB: 3 Gaps: 0

US-10-601-319-2 (1-440) x US-09-866-379-9 (1-1901)

QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
DB 188 ATGAAGCGGATCTTAATACCCATTTTATCTCTCTGATTCCGTTAACCCTGCAATCTGCA 247

QY 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
DB 248 TTGGCTCAGAGTGGAGCGGAGCTGAGCTGGAAGTGGTGGTATTGTTCAGTGGTCATGGT 307
QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrrpPro 60
DB 308 GTGGGTGCTCCACACAGGCCACGCACTGATGAGGATGTACCCACAGACGCGATGGCCA 367
QY 61 ThrTrrpProValLysLeuGlyTrrpLeuThrProArgGlyGlyGluLeuIleAlaTrrpLeu 80
DB 368 ACCTTGGCCGGTAAACTGGTGGCTGACACCGGNGGTGGTGAAGTAAATCGCCTATCTC 427
QY 81 GlyHisTrrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
DB 428 GGACATTACCAACGCCAGCGTCTGTAGCCGACGGATTGCTGGCGAAGAAAGGCTGCCCG 487
QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
DB 488 CAGTCTGGTCAAGTCCGGATTATTGCTGATGTCAGCAGCGGTACCCGTTAAACAGCGGAA 547
QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
DB 548 GCCTTCGCCCGCGGCTGGCACCTGACTGTGCAATAACCGTACATACCAGCAGATACG 607
QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
DB 608 TCCAGTCCCGATCCGTTATTAACTCTAANAACCTGGCGTTTGGCAACTGGATNAACGGG 667
QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
DB 668 AACGTGACTGACGGGATCTCTCAGCAGCGAGGAGGTCAATTGCTGACTTTACCGGGCAT 727
QY 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
DB 728 CGGCAAAACGGCGTTTCGGAACCTGGAACGGGTCTTAATTTTCGCAATCAAACTTTGTC 787
QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
DB 788 CTTAAACGTGAGAAACAGGACGAAAGCTGTTTCAATTAACGCGAGGCAATTACCATCGGAATC 847
QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
DB 848 AAGGTGAGCGCGCAATGCTCATTAACCGGTGCGGTAAAGCCCTCGCATCAATGCTGACG 907
QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrrpGlyArgIleThr 260
DB 908 GAGATATTTCTCTGCAACAGCAAGCAAGGGAATGCCGAGCGGGGTGGGGAAGATCAC 967
QY 261 AspSerHisGlnTrrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrrpLeuGln 280
DB 968 GATTCAACCCAGTGGAAACACTTGTAAAGTTTGCATTAACGGCAATTTTGTTCGTACAA 1027
QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
DB 1028 CGCACGCGCAGAGGTGGCCCGCAGCGGCCACCCCGTTATTAGATTTTGTATCAAGACAGCG 1087
QY 301 LeuThrProHisProProGlnLysGlnAlaTrrpGlyValThrLeuProThrSerValLeu 320
DB 1088 TTGACGCCCCCATCCACCGCAAAACAGGCGGTATGGTGTGACATTTACCACATTCAGTGTG 1147
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrrp 340
DB 1148 TTTATCGCCGGACACGATACTTAATCTGGCAAAATCTCGCGCGGCGCACTGGAGCTCACTGG 1207
QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrrp 360
DB 1208 ACGCTTCCCGGTGAGCGGATAACAGCCCGCAGGTGGTGAATCTGGTGTGTTGAACGCTGG 1267
QY 361 ArgArgLeuSerAspAsnSerGlnTrrpIleGlnValSerLeuValPheGlnThrLeuGln 380
DB 1268 CGTCGGCTAAGCGGATAACAGCCAGTGGATTTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1327
QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400

Db 1328 CAGATGGTGTAAACCCGCTGTCTAATTAATACGCCCGCCGAGAGGTGAACCTGACC 1387
 Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 Db 1388 CTGGCAGGATGTGAAGAGCGAATATGCGCAGGCATGTGTTGTTGGCAGGTTTACGCAA 1447
 Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
 Db 1448 ATCGTGAATGAAGCAGCATACCGCGTGCAGTTTG 1483

RESULT 8

US-09-540-149A-9
 ; Sequence 9, Application US/09540149A
 ; Patent No. 6511699
 ; GENERAL INFORMATION:
 ; APPLICANT: Lei, Xingen
 ; TITLE OF INVENTION: ENZYMES WITH IMPROVED PHYTASE ACTIVITY
 ; FILE REFERENCE: 19603/2791
 ; CURRENT APPLICATION NUMBER: US/09/540,149A
 ; CURRENT FILING DATE: 2000-03-31
 ; PRIOR APPLICATION NUMBER: 60/127,032
 ; PRIOR FILING DATE: 1999-03-31
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 9
 ; LENGTH: 1489
 ; TYPE: DNA
 ; ORGANISM: Escherichia coli
 US-09-540-149A-9

Alignment Scores:
 Pred. No.: 4,81e-225 Length: 1489
 Score: 2221.00 Matches: 428
 Percent Similarity: 99.1% Conservative: 0
 Best Local Similarity: 99.1% Mismatches: 4
 Query Match: 96.5% Indels: 0
 DB: 3 Gaps: 0

US-10-601-319-2 (1-440) x US-09-540-149A-9 (1-1489)

Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 Db 182 ATGAAACGATCTTAATCCCAATTTTATCTTTTATGATTGCTTAACCCCGCATCTGCA 241
 Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
 Db 242 TTGCTCAGAGTGAGCGGAGCTGAAGCTGMAAGTGTGGTGAATTGTTCAGCCGTCATGGT 301
 Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTTPPro 60
 Db 302 GTGCGTCCCCAACCAAGGCCACGCACTGATGAGGATGTACCCCGAGACGCGATGGCCA 361
 Qy 61 ThrTrpProValLysLeuGlyTTPLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
 Db 362 ACCTGGCCGGTAAACCTGGTGTGGCTGACACCGCGGTGGTGGCTAATCGCTATCTC 421
 Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 Db 422 GGACATTACCAACCCAGCGCTCTGTGGCCGACGAGATTGCTGGCGAAAGGGGTGCGCG 481
 Qy 101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 Db 482 CAGCTCGTCAAGTTCGGCTGATGTGATGTCACGAGCGTATCCCGTAAACAGGCGAA 541
 Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 Db 542 GCCTTCGCGCGCGGCTGGACCTGATGTGCAATAACCGTACATACCCAGGCGATACG 601
 Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 Db 602 TCCAGTCCCGATCGTTATTATTAATCTTAAATACTGGCGTTTGCCTAACTGGATAACGCG 661

Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 Db 662 AACGTCACTGACGCGATCTCTCAGCAGGCGAGGAGGTCAATTGCTGACTTTTACCGGCAT 721
 Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
 Db 722 CGGCAAAACGGCGTTTCGCGAACTGGAAACGGGTCTTAATTTTCCCAATTTAAACTGTGC 781
 Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 Db 782 CTTAACCGTGAAGAACAGGACGAAAGCTGTTCATTAAACGCGGAGCATTTACCATCGCACTC 841
 Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 Db 842 AAGTGAGCGCCGACAAATGTTTCAATACCGGTGGGTAGCTTCGATCAATGCTGACG 901
 Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
 Db 902 GAAATATTTCTCTGCAACAGCACAGGGAATCCCGGAGCCGGGTGGGAGAGATCACT 961
 Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
 Db 962 GATTCACACGAGTGAACACCTTGTCTAAGTTTGCATAACGCGCAATTTTATTACTACAA 1021
 Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 Db 1022 CGCACGCCAGAGGTTCGCGCAGTCGCGCCACCCGCTTATTTGATTTGATCATGGCAGCG 1081
 Qy 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
 Db 1082 TTGACGCCCATTCACCGCAAAACAGGCGTATGGTGTGACATTACCCACTTCCAGTGTG 1141
 Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 Db 1142 TTTATTGCGGACACGATATACTAATCTGGCAAAATCTCGCGCGCAGCTGGAGCTCACTGG 1201
 Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 Db 1202 ACCTTTCAGGTGAGCGCGGATACCGCCAGGTGGTGAATGTTGTTTGAACGCTGG 1261
 Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 Db 1262 CGTCGCTAAGCGATAACAGCCAGCTGGATTCAGGTTTCGCTGCTTCCAGACTTTACAG 1321
 Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 Db 1322 CAGATCGTGATAAACGCGCTATCATTTAAATACGCGCCGAGAGGTGAACACTGACC 1381
 Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 Db 1382 CTGGCAGGATGTGAAGCGAATATGCGAGGCGATGTGTTGTTGCGCGGTTTTTACGCAA 1441
 Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
 Db 1442 ATCGTGAATGAAGCGCGCATACCGCGTGCAGTTTG 1477

RESULT 9

US-10-266-041A-9
 ; Sequence 9, Application US/10266041A
 ; Patent No. 6974690
 ; GENERAL INFORMATION:
 ; APPLICANT: Lei, Xingen
 ; TITLE OF INVENTION: PHOSPHATASES WITH IMPROVED PHYTASE ACTIVITY
 ; FILE REFERENCE: 19603/2798
 ; CURRENT APPLICATION NUMBER: US/10/266,041A
 ; CURRENT FILING DATE: 2002-10-07
 ; PRIOR APPLICATION NUMBER: 60/127,032
 ; PRIOR FILING DATE: 1999-03-31
 ; PRIOR APPLICATION NUMBER: 09/540,149
 ; PRIOR FILING DATE: 2000-03-31
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 9

```
; LENGTH: 1489
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-266-041A-9

Alignment Scores:
Pred. No.: 4,81e-225 Length: 1489
Score: 2221.00 Matches: 428
Percent Similarity: 99.1% Conservatives: 0
Best Local Similarity: 99.1% Mismatches: 4
Query Match: 96.5% Indels: 0
DB: Gaps: 0

US-10-601-319-2 (1-440) x US-10-266-041A-9 (1-1489)

Qy 1 MetLysAlaIleLeuLeuProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
Db 182 ATGAAGCGATCTTAATCCATTTATCTCTTTGATTCCTTAACCCCGCAATCTGCA 241
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40
Db 242 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAAGTGTGTGATTTGTTCAGCCGTCATGGT 301
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 302 GTGCGTGCCCAACCAAGGCCACGCAACTGATGCAGAGATGTCAACCCAGACGATGGCCA 361
Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
Db 362 ACCTGGCCGGTAAATCTGGTTGGCTGCACACCACCGGGTGGAGCTAATCGCCCTATCTC 421
Qy 81 GlyHisTyrglnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 422 GGACATTACCAACGCGAGCGTGTGGCGGACGATTTGCTGGCGAAAGGGCTGCCCG 481
Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 482 CAGCCTGTGTGAGTCCGATATTATCTGATGTGAGCGGTACCGGTAAACACAGCGCAA 541
Qy 121 AlaPheAlaAlaGlyLeuAlaProPheCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 542 GCCTTCGGCGCGGGCTGGCCTGACTGTGCAATAACCGTATACCCAGGACGATACG 601
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAla 160
Db 602 TCCAGTCCGATCCGTTATTAATCTCTAANAACCTGGGTTGGCACTGGATACGCG 661
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db 662 AACGTGACTGACGCGATCTCAGCAGGCGAGGAGGTCAATTGCTGACTTTTACCGGCAT 721
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 722 CGGCAACCGCGGTTTCGGGAACCTGAACGGGGTCTTAATTTTCCCAATTAACCTGTGTC 781
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 782 CTTAACCTGTGAACACAGACGAAGCTGTTCATTAAACGAGGCAATTAACATCGGAATC 841
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 842 AAGGTGAGCGCGCAATGTTTCATTAAACGGTGGTAAGCCTCGCATCAATGCTGAGC 901
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Db 902 GAAATATTCTCTCTCAACAAGCAGGGAATGCGGAGCGGGGTGGGGAAGGATCACT 961
Qy 261 AspSerHisGlnTrpAsnThrLeuSerLeuHisAsnAlaGlnPheTyrlLeuLeuGln 280
Db 962 GATTCACACCACTGGGAACACCTTGTAGTTTGCATTAACGGCAATTTATTTACTCAA 1021
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 1021
```

```
Db 1022 CGCACGCCAGAGGTTGCCCGCAGTCGCCACCCCGTTATTGGATTTCATGATGCAGCG 1081
Qy 301 LeuThrProHisProGlnLysGlnAlaTyrglyValThrLeuProThrSerValLeu 320
Db 1082 TTGACGCCCATCCACCCGCAAAACAGCGGTATGTTGACATTAACCCACTTCAGTGTG 1141
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 1142 TTTATTGCGGACACGATACTAATCTGGCAAAATCTCGCGCGGCGACTGGAGCTCAACTG 1201
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlyLeuValPheGluArgTrp 360
Db 1202 ACGCTTCAGGTGACGCGGATAACACGCGCGAGGTGGTGAACCTGTTGTAACGCTGG 1261
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1262 CGTCGGCTTAACGATTAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCAGACTTTTACAG 1321
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyGluValLysLeuThr 400
Db 1322 CAGATGCGTGATAAAACGCCGCTATCATTAATAACGCCCGCGAGAGGTGAACCTGACC 1381
Qy 401 LeuAlaGlyCysGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1382 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGCGCATGTGTTGTTGGCGGTTTACGCAA 1441
Qy 421 IleValAsnGluAlaAatGlyProAlaCysSerLeu 432
Db 1442 ATCGTGAATGAAGCGCGCATACCGCGGTGCAGTTTG 1477

RESULT 10
US-09-866-379-6
; Sequence 6, Application US/09866379
; Patent No.: 6855365
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-6

Alignment Scores:
Pred. No.: 7,09e-225 Length: 1901
Score: 2221.00 Matches: 428
Percent Similarity: 99.1% Conservatives: 0
Best Local Similarity: 99.1% Mismatches: 4
```

```
Query Match: 96.5% Indels: 0
DB: 3 Gaps: 0
US-10-601-319-2 (1-440) x US-09-866-379-6 (1-1901)

Qy 1 MetLysAlaIleLeuProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
Db 188 ATGAAAGCGATCTTAATCCATTTTATCTCTTGATTCGGTTAAACCCGCAATCTGCA 247

Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40
Db 248 TTCCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGTGATGTGTCAGTCTGTCATGGT 307

Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
Db 308 GTGGTGCTCCACCAAGCCACCAACTGATGACGAGATGTCACCCAGACGCAATGGCCA 367

Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db 368 ACCTGGCCGGTAAAACTGGGTGGCTGACACCGCGNGGTGGTGAGCTAATCGCCTATCTC 427

Qy 81 GlyHisThrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 428 GGACATTACCTGGGCCAGCGCTCGGTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG 487

Qy 101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 488 CAGTCTGTCAGTTCGGATTAATGCTGATGTGACGAGCGTACCCGTTAAACAGGCGAA 547

Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCCTTCGCCGCGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 607

Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 608 TCCAGTCCCGATCCGTTATTTAATCCCTTAAAAACTGGCGTTGCCCACTGGATAACGCG 667

Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db 668 AACGTGACTGACGGATCCTCAGCAGGCGAGGGGTCAATTGCTGACTTTTACCGGGCAT 727

Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 728 CGCAAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTCCGCAATCAAACTTGTCG 787

Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 788 CTTAAACGTGAGAAACACAGGACGAAAGCTGTTCAATTAACGCGAGGCATTACCATCGAACTC 847

Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 848 AAGGTAGCCGCGCAATGCTCATTAACCGTGGTAAAGCTCGGATCAATGCTGACG 907

Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Db 908 GAGATATTTCTCTGCAACACAGCACAGGAAATGTCGAGCGGGGTGGGAGGATCACC 967

Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
Db 968 GATTACACACCATGGAAACACCTTGCTAAGTTTGCATTAACGCGCAATTTTATTTGCTACA 1027

Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 1028 CGCACCGCAGAGGTGGCCGAGCGCGCCACCCCGTTATTAGATTTTATGATCAAGACAGCG 1087

Qy 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 1088 TTGACGCCCATCCACCGCAAAAACAGCGGTATGTTGTGACATTTACCACCTTCAGTGCTG 1147

Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 1148 TTTATCGCCGACACGATCTAATCTGGCAAAATCTCGCGCGCGCACTGGAGCTCAACTGG 1207

Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 1208 ACCTTCCCGGTACGCGGATTAACACGCCGCCAGGTGGTGAACCTGGTTTGAACGCTGG 1267

Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1268 CGTCGGCTAAGCGATTAACACGCCAGTGATTCAGGTTTCGTGGTCTTCCAGACTTTACAG 1327

Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 1328 CAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCGAGAGGTGAACACTGACC 1387

Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1388 CTGGCAGGATGTGAAGCGAATATGCCAGGATGTGTTCTGTTGGCAGGTTTACGCA 1447

Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1448 ATCGTGAATGAAGCAGCATACATCCGCGTGCAGTTTG 1483

RESULT 11
US-09-866-379-5
; Sequence 5, Application US/09866379
; Patent No. 6855365
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866, 379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-5

Alignment Scores:
Pred. No.: 9,05e-225 Length: 1901
Score: 2220.00 Matches: 429
Percent Similarity: 99.3% Conservative: 0
Best Local Similarity: 99.3% Mismatches: 3
Query Match: 96.4% Indels: 0
DB: 3 Gaps: 0
US-10-601-319-2 (1-440) x US-09-866-379-5 (1-1901)

Qy 1 MetLysAlaIleLeuProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
Db 188 ATGAAAGCGATCTTAATCCATTTTATCTCTTGATTCGGTTAAACCCGCAATCTGCA 247

Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40
```

Db 248 TTCCGTCAGAGTACCGCGAGCTGAAGCTGGAAAGTGTGGTGTGATTTGTCTAGTCTCATGGT 307
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 308 GTGCGTGTCTCCAACAGCCACGCACTGATGCAGGATGTCAACCCAGAGCATGGGCA 367
Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db 368 ACCTGGCCGGTAAACTGGTTGACTGACACCGCGNGGTGGTGAGCTAATCGCCTATCTC 427
Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
Db 428 GGACATTACCAACGCGCAGCTGTGTAGCCGAGATTGCTGGCGAAAGGGCTGCCCG 487
Qy 101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 488 CAGCTGTGTAGTCCGATTAATCTGATGTCGACGAGCGGTACCCGTAACAGCGCAA 547
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCCTTCGCGCGCGGCTGCACCTGACTGTGCAATAACCGTACATACCAGGAGATACG 607
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAla 160
Db 608 TCCAGTCCCGATCCGTTAATCTCTAATAAAGCTGGCGTTTGCAACTGGATAACGCG 667
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrClyHis 180
Db 668 AACGTGACTGACGCGATCTCAGCAGGCGAGGAGGTCAATTGCTGACTTTACCGGGCAT 727
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 728 CGGCAACCGCGTTTCGGGAACCTGGAACGGGTGCTTAATTTCCGCATCAAACTTGTGC 787
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 788 CTTAACGCTGGAACACGAGCAAGCTGTTCATTAACGCGAGCATTAACCATCGGAATC 847
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 848 AAGGTGAGCGCGACATGTCTCATTAACCGGTGCGTAAGCCTCGCATCAATGCTGACG 907
Qy 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Db 908 GAGATATTTCTCTGCAACAGCACAGGAGATGCGGAGCGCGGGTGGGGAGGATCAC 967
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 968 GATTACACACAGTGGAAACACCTTGCTAAGTTTGCAATACGCGCAATTTAATTTGCTACA 1027
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 1028 CGCAGCCAGAGGTTGCGCGAGCGCGCCACCCCGCTTATTAGATTTGATCAAGACAGCG 1087
Qy 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 1088 TTGAGCCCCCATCCACCCGAAACAGCGGTATGTTGTGTGACATTAACCATTCAGTGTG 1147
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 1148 TTTATCGCGGACACAGATACTAATCTGGCAATCTCGCGCGGCGACTGGAGCTCAACTGG 1207
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlyLeuValPheGluArgTrp 360
Db 1208 ACGCTTCGCGTCAGCCGGATAACAGCGCGAGGTGGTGAAGTGTGCTGCTTCCAGACTTAC 1267
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1268 GCTCGGCTAAGCGATAACAGCCAGGTGGATTGAGGTTTCGCTGCTTCCAGACTTACAG 1327
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400

Db 1328 CAGATGCGTGATAAAACCGCCGCTCTGCTAATAATACGCGCCCGGAGAGGTGAACCTGACC 1387
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1388 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGSCATGTGTTCTGTGGCAGGTTTACGCAA 1447
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1448 ATCGTGAATGAACGACGACATACCGCGGTGCAGTTTG 1483
RESULT 12
US-09-715-477-2
; Sequence 2, Application US/09715477
; Patent No. 6841370
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYTASE
; FILE REFERENCE: 19603/4031
; CURRENT APPLICATION NUMBER: US/09/715.477
; CURRENT FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,179
; PRIOR FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1486
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-715-477-2
Alignment Scores:
Pred. No.: 1,27e-224 Length: 1486
Score: 2217.00 Matches: 427
Percent Similarity: 98.8% Conservative: 0
Best Local Similarity: 98.8% Mismatches: 5
Query Match: 96.3% Indels: 0
DB: Gaps: 0
US-10-601-319-2 (1-440) x US-09-715-477-2 (1-1486)
Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 188 ATGAAGCGCATCTTAATCCCATTTTATCTCTCTGATTCCGTTTAACCCGCAATCTGCA 247
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40
Db 248 TTCGCTCAGGTGAGCGGAGCTGAAGCTGGAAGTGTGGTATTGTGAGCCGTATGCT 307
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 308 GTGCGTCCCCCAACCAAGGCCACCACTGATGACGAGATGTCAACCCAGAGCATGGCA 367
Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db 368 ACCTGGCCGGTAAACTGGGTGGCTGACACACGCGGTGGTGAGCTAATCGCCTATCTC 427
Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
Db 428 GGACATTACCAACGCGCAGCGTCTGTGGCGGACGAGTGTGGCGGAAAGGGCTGCCCG 487
Qy 101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 488 CAGCCTGTGTAGTCCGATTAATTTGTCGATGTCCAGCGCGTACCCGTAAACAGCGCAA 547
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCCTTCGCGCGCGGCTGCACCTGACTGTGCAATAACCGTACATACCAGGAGATACG 607
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAla 160
Db 608 TCCAGTCCCGATCCGTTAATTTATTTCTCTAATAAAGCTGGCGTTTGCCCACTGGATAACGCG 667
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180

```
Db 668 AACGTGACTAGCGGATCTCTACGACGGCGAGGGGTCAATTGCTGACTTTTACCGGCAT 727
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 728 CGGCAAAACGGCGTTTCGCGAATGGAACGGGTGCTTAATTTTCGCAATCAAACTTGTC 787
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 788 CTTAAACGTGAGAAACAGGACGAAAGCTGTTTCAATACGCGAGGCATTTACCATCGGAATC 847
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 848 AAGGTGAGCGCCGACAAATGTTTCAATACCGGTGCGTAAGCCTTCGATCAATGCTGACG 907
Qy 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Db 908 GAAATATTTCTCTCGCAACAGCAGGGAATGCCGAGCGCGGGTGGGGAAGATCACT 967
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
Db 968 GATTACACAGGTGGAACACCTTGTCTAAGTTTGCATTAACGCGCAATTTTATTACTACAA 1027
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 1028 CGCACGCCAGAGGTTCGCCGAGTTCGCCACCCCGTATTGGATTGGATTCAGACAGCG 1087
Qy 301 LeuThrProHisProGlnLysGlnAlaArgGlyValThrLeuProThrSerValLeu 320
Db 1088 TTGACGCCCATCCACCGCAAAACAGCGGTATGGTGTGACATTTACCCACTTCAGTGCTG 1147
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp 340
Db 1148 TTTATTCCCGGACACGATACTAATTCGCGAAATCTCGCGCGCGCACTGGAGCTCAACTGG 1207
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 1208 ACGTTCCAGTCCGCGGATTAACACGCCCGCAGGTGGTGAATGGTGTGAACTGAACTGG 1267
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1268 CGTCGGCTAAGCGATACACGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1327
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 1328 CAGATGGCTGATAAAACCGCGCTATCATTAATAACGCCCGCGAGAGGTGAAACTGACC 1387
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1388 CTGGCAGGATGTGAGAGCGGAAATGCCAGGCAATGTGTTCTGTTGGCCGTTTACGCAA 1447
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1448 ATCGTGAATGAAGCGCGCATACCGCGCTGCAGTTTG 1483
```

RESULT 13

```
US-09-715-477-4
; Sequence 4, Application US/09715477
; Patent No. 6841370
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYTASE
; FILE REFERENCE: 19603/4031
; CURRENT APPLICATION NUMBER: US/09/715,477
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,179
; PRIOR FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1486
; TYPE: DNA
; ORGANISM: Escherichia coli
```

```
US-09-715-477-4
Alignment Scores:
Pred. No.: 1,67e-222 Length: 1486
Score: 2197.00 Matches: 424
Percent Similarity: 98.6% Conservative: 2
Best Local Similarity: 98.1% Mismatches: 6
Query Match: 95.4% Indels: 0
DB: 3 Gaps: 0

US-10-601-319-2 (1-440) x US-09-715-477-4 (1-1486)
Qy 1 MetLysAlaIleLeuLeuProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 188 ATGAAGAGCATCTTAATCCCATTTTATCTTCTGATTCCTGTAATACCCGCAATCTGCA 247
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db 248 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGGTGAATTTGTACGCCGTCACTGT 307
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 308 GTGCGTGGCCCAACCAAGGCCACGCAACTGATGCAGGATGTCAACCCAGACGATGSCCA 367
Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
Db 368 ACCTGCGCGGTAAACCTGGGTGGCTGCACCAACGCGGTGGTGAATTCGCTTATCTC 427
Qy 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 428 GGAATATTACCAACGCCAGCGCTCTGGTGGCCGACGATTCGTGGCAAAAAGGCGTCCCG 487
Qy 101 GlnSerGlyGlnValAlaIleLeuAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 488 CAGCTTGGTTCAGGTTCGATTAATTCGATGTGCAGCAGCGTACCGGTAAACACAGCGCAA 547
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCCTTCGCGCGCGGTGGCACCTGACTGTGCAATTAACCGTACATACCCAGGAGATACG 607
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 608 TCCAGTCCCGATCCGTTATTATTCTCTAANAACCTGGCGTTTCCCACTGGATACGCG 667
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db 668 AACGTGACTGACGCGATCTCTCAGCAGGCGAGGAGGTCAATTGCTGACTTTTACCGGCAT 727
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 728 CGGCAAAACGGCGTTTCGCGAATCGGAACGCGGTCTTAATTTTCCCAATCAAACTTGAAC 787
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 788 CTTAAACGTGAGAAACAGAAATGAAGCTGTAACTTAACCGCGGTGAAGCCTTCGATCAATGCTGACG 847
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 848 AAGGTGAGCGCCGACAAATGTTTCAATTAACCGGTGCGTAAGCCTTCGATCAATGCTGACG 907
Qy 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Db 908 GAAATATTTCTCTCGCAACAGCAGGGAATGCCGAGCGCGGGTGGGGAAGATCACT 967
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
Db 968 GATTACACAGGTGGAACACCTTGTCTAAGTTTGCATTAACGCGCAATTTTATTACTACAA 1027
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 1028 CGCACGCCAGAGGTTCGCCGAGTTCGCCACCCCGTATTGGATTGGATTCAGACAGCG 1087
Qy 301 LeuThrProHisProGlnLysGlnAlaArgGlyValThrLeuProThrSerValLeu 320
```

Db 1088 TTGACGCCCATCCACCGCAAAACAGGCGTATGGTGTGACATTACCCACATTCAGTGCTG 1147
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyValAlaLeuGluLeuAsnTrp 340
Db 1148 TTTATTGCGGACACAGTACTAATCTGGCAATCTCGGCGGCGCACTGGAGCTCAACTGG 1207
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 1208 AGCTTCCAGGTCAGCGCGGATAACACGCGCCAGGTGGTGAATCTGGTGTGTAACGCTGG 1267
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpLeGlnValSerLeuValPheGlnThrLeuGln 380
Db 1268 CGTCGGCTAAGCGGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1327
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 1328 CAGATGCGTGATTAACCGCGCTATCTAATAATACGCGCGCGAGAGGTGAATCTGACC 1387
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1388 CTGGCAGGATGTGAAGAGCGAAATCGCGAGGCGCATGTGTTCGTTGGCCGGTTTACGCA 1447
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1448 ATCGTGAATGAAGCGCGCATACCGCGGTGCGATTTG 1483

RESULT 14

US-08-910-798-1
; Sequence 1, Application US/08910798
; Patent No. 5876997
; GENERAL INFORMATION:
; APPLICANT: KRETZ
; TITLE OF INVENTION: NOVEL PHYTASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: August 13, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HAILLE, PH.D., LISA A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09010/029001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1272 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: PHYTASE
; FEATURE:
; NAME/KEY:
; LOCATION:
US-08-910-798-1
Alignment Scores:

Pred. No.: 2,39e-222 Length: 1272
Score: 2194.50 Matches: 423
Percent Similarity: 96.1% Conservative: 0
Best Local Similarity: 96.1% Mismatches: 0
Query Match: 95.3% Indels: 17
DB: 2 Gaps: 1
US-10-601-319-2 (1-440) x US-08-910-798-1 (1-1272)
Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAGCGATCTTAATCCCATTTTATCTCTCTGATTCCGTAAACCCCGCAATCTGCA 60
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db 61 TTCGCTCAGATGAGCGGAGCTGAGCTGAAAGTGGTGGTATTGTGATGTCATGTCATGT 120
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 121 GTGCGTGTCTCAACCAAGGCCACGCAACTGATGCAGGATGTCAACCCAGAGCGCATGGCCA 180
Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
Db 181 ACCTGGCGCGTAAAACTGGGTGGCTGACACCGCGNGGTGGTGAAGCTAATCGCCTATCTC 240
Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 241 GGACATTACCAACGCCAGCGTCTGGTAGCCGAGATTTGTCGCCGAAAGAGGCTGCCCG 300
Qy 101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 301 CAGTCTGCTCAGGTGCGGATTTATTTGCTGATGTCGACGAGCGTACCCGTAAACAGGC--- 357
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 358 -----CAGCGCATACG 369
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 370 TCCAGTCCGATCCGTTATTTAATCTCTAANAACCTGGCGTTTGCACACTCGATNACCG 429
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db 430 AACGTGACTGACGCGATCTCTCAGCAGGCGAGGCGTCAATTTGCTGACTTTTACCGGCAT 489
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 490 CGGCANAACGGGTTTCGCGAACTGGAAACGGGTGCTTAATTTTCGCGCAATCAAACTTGTGC 549
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 550 CTTAAGCTGNAANAACAGGACGAAGCTGTTTATTAACGAGGATTTACCATCGGAATC 609
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 610 AAGGTGAGCGCGCAATGCTCTATTAAACCGTGGTGGTGAAGCTTCGATCAATCTGACG 669
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Db 670 GAGATATTCTCTGCAACAAGCACAGGGAATGTCGAGGCGCGGGGTGGGAAGGATCAC 729
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
Db 730 GATTACACCACTGGGAACACTTGTGAAGTTTGCATTAACGGGCANATTTTATTGCTACAA 789
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 790 CGCAGCGCAGAGGTGCGCGCAGCGCCACCCCGTTATTGGATTTGATCATCGCAGCG 849
Qy 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 850 TTGACGCCCATCCACCGCAAAACAGGCGGTATGGTGTGATGATTAACCATTCAGTACTG 909

QY 321 PheIleAlaGlyHisAspThrAenLeuAlaAsnLeuGlyAlaLeuLeuLeuAenTyr 340
 Db 910 TTTATTGCGGACACGATCTAATCTGCGCAATCTCGCGCGCACCTGAGAGCTCAACTGG 969
 QY 341 ThrLeuProGlyGlnProAspAenThrProProGlyGlyGluLeuValPheGluAatTyr 360
 Db 970 ACGTTCCCGGTGACGCGGATACACGCGCGCAGGTGGTGAATGGTGTGTTGAACGCTGG 1029
 QY 361 ArgArgLeuSerAspAenSerGlnTyrPileGlnValSerLeuValPheGlnThrLeuGln 380
 Db 1030 CGTCGGCTAAGCGATAACAGCCAGTGTATCAGGTTTCGGTGGTCTTCCAGACTTTACAG 1089
 QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 Db 1090 CAGATGCGTGATAAAGCGCGCTGTCTAATAATACGCGCGCGGAGAGGTGAACACTGACC 1149
 QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 Db 1150 CTGGCAGATGTGAGAGCGAATGCGCAGGGCATGTGTTCTGGTGGCAGGTTTACGCAA 1209
 QY 421 IleValAenGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
 Db 1210 ATCGTGAATGAAGCAGCATACCGCGCTGCAGTTTGAGATCTCATCACCATCACCATCAC 1269

RESULT 15
 US-09-489-039A-341
 ; Sequence 341, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLBBSIELLA
 ; FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 341
 ; LENGTH: 1266
 ; TYPE: DNA
 ; ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-341

Alignment Scores:
 Pred. No.: 6.89e-46 Length: 1266
 Score: 528.00 Matches: 134
 Percent Similarity: 48.7% Conservative: 69
 Best Local Similarity: 32.1% Mismatches: 184
 Query Match: 22.9% Indels: 30
 DB: 3 Gaps: 10

US-10-601-319-2 (1-440) x US-09-489-039A-341 (1-1266)

QY 19 SerAlaPheAlaGlnSerGlu-----ProGlu---LeuLysLeuGluSerVal 33
 Db 67 TCCGCTGGCGCTCAGGCGCGACGACCAAGCGCGCCCGAGGGGTATCAGCTGCACAGGTG 126
 QY 34 ValIleValSerArgHisGlyValArgAlaPro---ThrLysAlaThrGlnLeuMetGln 52
 Db 127 CTGATCATGAGCCCTCACACCTCGTGGTGGCGCGTGGCCCAATAACGGCAGCGTCTGGAA 186
 QY 53 AspValThrProAspAlaTyrProThrTyrProValLysLeuGlyTyrLeuThrProArg 72
 Db 187 CAGTCCACCGCAAGGCTGCGCGCAGTGGGATGTCGCGCGCGCCGCGCAGCTGACCACCAA 246
 QY 73 GlyGlyGluLeuIleAlaTyrLeuGlyHisTyrGlnArgGlnArgLeuValAlaAspGly 92
 Db 247 GCGCGTGTCTGGAGGTGTATATGGCCACTATATGCGCGGAATGGCTGGCGCCACGAGAAG 306
 QY 93 LeuLeuAlaLysLysGlyCysProGlnSerGlyGlnValAlaIleAlaAspValAsp 112
 Db 307 CTGGTGACCGCGCGAGTGGCCCGCGGAAACCGCGTTTATGCTTACGCTAACACGCTG 366

Search completed: June 14, 2006, 15:34:29

Job time : 252.541 secs

QY 113 GluArgThrArgLysThrGlyGluAlaPheAlaAlaGlyLeuAlaProAspCysAlaIle 132
 Db 367 CAGCGACCGCTCCGCCACCGCGAGTTCTTTATCACCAGCGCGTTCGCGGCTCGCGGATC 426
 QY 133 ThrValHisThrGlnAlaAspThrSerSerProAspProLeuPheAenProLeuLysThr 152
 Db 427 CCGGTGATCACCAGCCTCAGATGGGCACCATGACCCCGACCTTCAACCCGTTATTATACC 486
 QY 153 GlyValCysGlnLeuAspAsnAlaAsnValThrAspAlaIle---LeuSerArgAlaGly 171
 Db 487 GACGACTCCCGCGGTTTCGCGAGAAAGCGCTGCAGGCGATGGAAAAAGAGCGCGGGA 546
 QY 172 GlySerIleAlaAspPheThrGlyHisArgGlnThrAlaPheArgGluLeuGluArgVal 191
 Db 547 ATGCAAGTTGACCGAG-----AGCTACAAGCTGCTGGAGACCATG 585
 QY 192 LeuAsnPheProGlnSerAsnLeuCysLeuLysArgGluLysGlnAspGluSerCysSer 211
 Db 586 ATTGACTACCGCAACTCGCCCTCTGCTG-----AAGAGAGAGAGGTCTGTTC 633
 QY 212 LeuThrGlnAlaLeuProSerGluLeuLysValSerAlaAspAsnValSerLeuThrGly 231
 Db 634 CTGAGCGAGGTAAAGATACCTTTAGCGCGGTTATCAGCAGGAGCGCGGAGTGTCCGA 693
 QY 232 AlaValSerLeuAlaSerMetLeuThrGluIlePheLeuLeuGlnAlaGlnGlyMet 251
 Db 694 CCGTTGAAGGTGGGCAACTCGCTGGTAGACGCTTCACCTGCAATATTATGAAGGTTC 753
 QY 252 Pro-----GluProGlyTyrGlyArgIleThrAspSerHisGlnTyrAsnThrLeuLeu 269
 Db 754 CCAGAGATCAGTGGCGTGGGGAGATGCCAGCGATGAAGCATGGCGGAGTCTGTCTCG 813
 QY 270 SerLeuHisAsnAlaGlnPheTyrLeuLeuGlnArgThrProGluAlaAlaArgSerArg 289
 Db 814 AAGCTGAATAATGCTATCAGGATAGTCTGTTTACCTCGGTGGCGGTGGCGCAAAACGTC 873
 QY 290 AlaThrProLeuLeuAspLeuIleMetAlaAlaLeuThrProHisProGlnLysGln 309
 Db 874 GCCAAACCGCTGTTAAATATATATGATACGCGCTGTTGGCGCGGAGCGGAGCGCAAG--- 930
 QY 310 AlaTyrGlyValThrLeuProThrSerValLeuPheIleAlaGlyHisAspThrAsnLeu 329
 Db 931 -----GCGAAGTCACTGCTGCTGGTGGGACACGACTCGAATATC 969
 QY 330 AlaAsnLeuGlyGlyAlaLeuGluLeuAsn---TrpThrLeuProGlyGlnProAspAsn 348
 Db 970 GCTTCGCTGCTCACCGCCCTGGACTTTAAGCCCTATCAACTGCGCGCGCAGTATGAGCGC 1029
 QY 349 ThrProProGlyGlyGluLeuValPheGluArgTyrArgArgLeuSerAspAsnSerGln 368
 Db 1030 ACTCCTATCGCGGCAAACTGCTGTTTCAGCGCTGGCATGCACGCGCGGCAATCGCGAT 1089
 QY 369 TrpIleGlnValSerLeuValPheGlnThrLeuGlnMetArgAspLysThrProLeu 388
 Db 1090 CTGATGAGATTGATGATGCTATCAAGTACCGAGCAACTGCGTAACCGCGCAGCGCTTA 1149
 QY 389 SerLeuAsnThrProProGlyGluValLysLeuThrLeuAlaGlyCysGluGluArgAsn 408
 Db 1150 ACCGTGCGGCGCGCGCGAGCGGTCACTCTGCGCTTAACCGCTGTCCGGTT---GAT 1206
 QY 409 AlaGlnGlyMetCysSerLeuAlaGlyPheThrGlnIleValAenGluAla 425
 Db 1207 GACGAGGGATTCTGCCGCTGGAGACCTTTAAAAATGGTGTATCAACGAGGCG 1257

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 14, 2006, 10:38:44 ; Search time 562.615 Seconds
(without alignments)
8179.104 Million cell updates/sec

Title: US-10-601-319-2

Perfect score: 2302

Sequence: 1 MKAILPPLSLILPLTQSA.....IVNEARIPACSLRSHHHHH 440

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/absz/ABSSWEB_spool/US10601319/runat_12062006_180047_27417/app_query.fasta_1
-DB=N_Geneseq -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=absz06h
-USER=US10601319 -CGEN 1 1 749 @runat_12062006_180047_27417 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq 8:*

1: Geneseq1980s: *
2: Geneseq1990s: *
3: Geneseq2000s: *
4: Geneseq2001as: *
5: Geneseq2001bs: *
6: Geneseq2002as: *
7: Geneseq2002bs: *
8: Geneseq2003as: *
9: Geneseq2003bs: *
10: Geneseq2003cs: *
11: Geneseq2003ds: *
12: Geneseq2004as: *
13: Geneseq2004bs: *
14: Geneseq2005s: *
15: Geneseq2006s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2302	100.0	1323	5 AAC88885	Aac88885 Escherich
2	2302	100.0	1323	6 AAD36473	Aad36473 Escherich
3	2302	100.0	1323	6 AAD25460	Aad25460 Escherich

4	2302	100.0	1323	8 ADA19445	Ada19445 E. coli B
5	2302	100.0	1323	12 ADO50291	Ado50291 Escherich
6	2302	100.0	1323	15 AEE75412	Aee75412 E. coli B
7	2282	99.1	1323	12 ADO50303	Ado50303 Escherich
8	2279	99.0	1323	12 ADO50301	Ado50301 Kangaroo
9	2235	97.1	1299	3 ABK12514	Abk12514 DNA encod
10	2235	97.1	1299	8 ACA19297	ACA19297 Prokaryot
11	2235	97.1	1299	12 ADL16137	Adl16137 Escherich
12	2235	97.1	1901	6 AAD25463	Aad25463 Escherich
13	2235	97.1	1901	8 ADA19449	Ada19449 E. coli K
14	2235	97.1	1901	10 ADC87744	Adc87744 DNA encod
15	2235	97.1	1901	12 ADO50297	Ado50297 Escherich
16	2235	97.1	1901	14 AED50819	Aed50819 Escherich
17	2235	97.1	1901	15 AEE75418	Aee75418 E. coli B
18	2235	97.1	3470	3 AAC68298	Aac68298 R15/APPA
19	2235	97.1	4060	3 AAC68296	Aac68296 R15/APPA
20	2235	97.1	6116	3 AAC68297	Aac68297 R15/APPA
21	2235	97.1	6708	3 AAC68295	Aac68295 R15/APPA
22	2235	97.1	17732	3 AAC68300	Aac68300 Lama2/APP
23	2235	97.1	20623	3 AAC68294	Aac68294 Lama2/APP
24	2230	96.9	1296	12 ADL16139	Adl16139 Shigella
25	2228	96.8	1901	12 ADO50296	Ado50296 Escherich
26	2224	96.6	1296	12 ADL16138	Adl16138 Shigella
27	2221	96.5	1489	3 AAC28216	Aac28216 E. coli a
28	2221	96.5	1489	8 ACC57672	Acc57672 Escherich
29	2221	96.5	1901	12 ADO50295	Ado50295 Escherich
30	2221	96.5	1901	15 AEE75417	Aee75417 E. coli p
31	2221	96.5	1901	15 AEE75416	Aee75416 E. coli p
32	2218	96.4	1315	10 ADL02197	Adl02197 DNA encod
33	2217	96.3	1486	4 AAD06831	Aad06831 E. coli a
34	2205	95.8	5421	3 AAC68299	Aac68299 SV40/APPA
35	2197	95.4	1486	4 AAD06832	Aad06832 E. coli a
36	2197	95.4	1486	8 ACC57673	Acc57673 Escherich
37	2194.5	95.3	1272	2 AAX26540	Aax26540 DNA encod
38	2188	95.0	1308	10 ADC87742	Adc87742 DNA encod
39	2188	95.0	1308	12 ADO50299	Ado50299 Escherich
40	2188	95.0	1308	15 AEE75420	Aee75420 E. coli B
41	2186	95.0	1299	13 ADW76362	Adw76362 Phytase A
42	2185	94.9	1299	15 AEF15218	Aef15218 Escherich
43	2171	94.3	1296	14 AED50817	Aed50817 Escherich
44	2171	94.3	1901	15 AEE75422	Aee75422 E. coli B
45	2147	93.3	1901	8 ADA19452	Ada19452 E. coli K

ALIGNMENTS

RESULT 1

AAC88885

ID AAC88885 standard; DNA; 1323 BP.

AC AAC88885;

XX

DT 07-MAR-2001 (first entry)

XX

DE Escherichia coli B phytase enzyme nucleotide sequence.

XX

KW Escherichia coli B; phytase enzyme; anabolic; phytate digestion;

XX nutrition; ds.

XX Escherichia coli.

OS

XX WO200071728-A1.

FN

XX 30-NOV-2000.

PD

XX 25-MAY-2000; 2000WO-US014846.

PF

XX 25-MAY-1999; 99US-00318528.

PR

XX (DIVE-) DIVERSA CORP.

PA

XX Short JM, Kretz KA;

XX

DR WPI; 2001-112081/12.
 DR P-PSDB; AAB37892.
 XX
 PT Improving the nutritional value of phytate-containing foodstuffs, using
 PT phytase enzymes which catalyze the liberation of inorganic phosphate from
 PT the phytates.
 XX
 PS Claim 2; Fig 1; 147pp; English.
 XX
 CC The present sequence encodes a phytase enzyme from *Escherichia coli* B.
 CC The enzyme catalyses the liberation of inorganic phosphate from the
 CC phytate in phytate-containing foodstuffs and can thus be used to improve
 CC the nutritional value of phytate rich ingredients
 XX
 SQ Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	3.15e-202	Length:	1323
Score:	2302.00	Matches:	440
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	5	Gaps:	0

US-10-601-319-2 (1-440) x AAC88885 (1-1323)

Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 Db 1 ATGAAGCGATCTTAATCCCATTTTATCTCTTCGATTCCGTTAAACCCGCAATCTGCA 60
 Qy 21 PheAlaGlnSerGluProGluLeuLeuLeuGluSerValValIleValSerArgGHisGly 40
 Db 61 TTCGCTCAGAGTAGCGAGCTGAAGCTGGAAGTGTGGTGTGATGTGCTGCTGTCATGGT 120
 Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 Db 121 GTGGTGCTCCACCAAGGCCACCGCAACTGATGTCAGGATGTCAACCCAGACGATGSCCA 180
 Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrlLeu 80
 Db 181 ACCTGGCCGGTAAACCTGGGTGGCTGCACACCGCGNGGTGGTGAGCTTAATCGCCTATCTC 240
 Qy 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaTyrlLysGlyCysPro 100
 Db 241 GGACATTTACCAACCCACGCGTCTGGTAGCCGACGGATGTGCGCAAAAAGGGCTGCCCG 300
 Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 Db 301 CAGTCTGGTCAGGTCCGGATTATTGCTGATGTGACGAGCGGTACCCGTAAACACAGGCGAA 360
 Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 Db 361 GCCTTCGCGCGCGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGCGAGATACG 420
 Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 Db 421 TCCAGTCCCGATCCGTTATTTAAATCCCTCAAAAACCTGGCGTTTCCCACTGGATAACGCG 480
 Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
 Db 481 AACGTGACTACCGGATCCTCAGCAGCGCAGGAGGGTCAATTGCTGACTTTACCGGSCAT 540
 Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
 Db 541 CGGCAAAACGGCGTTTCGCAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600
 Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 Db 601 CTTAAACGTGGAACACAGGACGAAGCTGTTCATTAAACGCGAGGATTTACCATCGGAATC 660
 Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 Db 661 AAGGTGAGCGCGACAAATGCTCTATTAAACCGGTGCGGTAAAGCCTCGCATCAATGCTGACG 720

Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
 Db 721 GAGATATTTCTCTCGCAACAGCACAGGGAATCCCGAGCCGGGTGGGGAAGGATCACC 780
 Qy 261 AspSerHisGlnTrpAsnThrLeuSerLeuHisAsnAlaGlnPheTyrlLeuLeuGln 280
 Db 781 GATTCAACACGTTGGAAACACCTTGTCTAAGTTTGCATTAACGCGCAATTTATTGCTACAA 840
 Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 Db 841 CGCACGCCAGAGGTGTCGCGCAGCCGCCACCCCGTATTGATTTGATCATGGCAGCG 900
 Qy 301 LeuThrProHisProGlnLysGlnAlaTyrlGlyValThrLeuProThrSerValLeu 320
 Db 901 TTGACGCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTACTG 960
 Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 Db 961 TTTATTGCGGACACGATACTAATCTGGCAAAATCTCGGCGCGCACTGGAGCTCACTGG 1020
 Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 Db 1021 ACCTTCCCGGTGAGCGGATAACCGCGCAGGTGGTGAACCTGGTGTGAAACGCTGG 1080
 Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 Db 1081 CGTCGCTTAAGCGATAACAGCCAGTGGATTTCAGGTTTCGCTGCTTCCAGACTTTACAG 1140
 Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 Db 1141 CAGTGGGTGATAAAGCGCGCTGTCTATTAAATACGCGCGCCGAGAGGTGAAACTGACC 1200
 Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 Db 1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGCGATGTGTTGCTGGCAGGTTTTACGCAA 1260
 Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
 Db 1261 ATCGTAATGAAGCACGATACCGCGGTGAGTTTGAGATCTCATCACCATCCATCAC 1320

RESULT 2
 AAD36473
 ID AAD36473 standard; DNA; 1323 BP.
 XX
 AC AAD36473;
 XX
 DT 21-AUG-2002 (first entry)
 XX
 DE *Escherichia coli* phytase DNA.
 XX
 KW Dietary aid; biocompatible composition; therapeutic; digestive tract;
 KW foodstuff; digestion; phytase; enzyme; gene; ds.
 XX
 OS *Escherichia coli*.
 XX
 FH Key Location/Qualifiers
 CDS 1..1323
 FT /*tag= a
 FT /product= "E. coli phytase protein"
 XX
 PN WO200189317-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 15-MAY-2001; 2001WO-US015764.
 XX
 PR 25-MAY-2000; 2000US-00580937.
 XX
 PA (DIVE-) DIVERSA CORP.
 XX
 PI Short JM, Kretz KA, O'donoghue E;
 XX

DR WPI; 2002-164149/21.
DR P-PSDB; AAE22836.

PT New dietary aids comprising sustained release biocompatible compositions,
PT comprise agent that assists in digestion, useful for delivering enzymes,
PT therapeutics, medicine or agents to an organism.

PS Disclosure; Fig 1; 89pp; English.

XX The present invention relates to novel dietary aids comprising sustained
CC release biocompatible composition which comprises an agent (enzymes such
CC as phytase, amylase, esterase, protease) that assists in digestion. The
CC biocompatible composition is effective upon oral consumption and release
CC in the digestive tract of a subject. The dietary aids are useful for
CC delivering enzymes, therapeutics, medicine and agents to an organism. The
CC use of enzymes and other agents in digestive aids of livestock or
CC domesticated animals not only improves the animal's health and life
CC expectancy but also assists in increasing the health of livestock or in
CC the production of foodstuffs from livestock. The present sequence is
CC Escherichia coli phytase DNA

SQ Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 3.15e-202 Length: 1323
Score: 2302.00 Matches: 440
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-601-319-2 (1-440) x AAD36473 (1-1323)

Qy 1 MetLysAlaLeuLeuProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
Db 1 ATGAAGCGATCTTAATCCCATTTTATCTCTTGTGATTCGTTAAACCCCGCAATCTGCA 60
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40
Db 61 TTCGCTCAGATGAGCGGAGCTGAAGCTGGAAAGTGTGGTGTGATTCAGTCGTATGTT 120
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 121 GTGGTGTCTCCAAACCAAGCCAGCACTGATGCGAGGATGTCAACCCAGAGCGCATGGCA 180
Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrlu 80
Db 181 ACCTGGCCGGTAAACCTGGGTGGCTGACACCGGNGGTGTGAGCTAATCGCCTATCTC 240
Qy 81 GlyHisTyrluGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 241 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGGATTGCTGGCGAAAAGGGTGGCCG 300
Qy 101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 301 CAGTCTGCTCAGGTCCGGATTAATGCTGATGTCACAGCGGTACCCGTAACAGCGGAA 360
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCCTTCGCGCGCGGGCTGGACCTGACGTGTCATTAACCGTACATACCCAGCGAGATACG 420
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCCAGTCCCGATCCGTATTATTATCCCTAANAACCTGGCGTTTCCCAACTGGATAACGG 480
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGATGACGGGATCCTCAGCAGGCGAGGGSTCAATGTGCTTACCGGGCAT 540
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 541 CGGCAACCGCGGTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAACTTGTGC 600

Qy 201 LeuLysArgGlyLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTAAACGCTGAGAAACAGACGAAAGCTGTTCAATTAACGAGGCATTACCATCGNACTC 660
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGAGCGCGCAATGTCTCATTAACCGGTGCGTAAGCCTCGCATCAATGTGTGAG 720
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnLysMetProGluProGlyTrpGlyArgIleThr 260
Db 721 GAGATATTTCCTCGCAACAGACAGGGAATGCCGAGCGCGGGTGGGAAAGGATCACC 780
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrluLeuGln 280
Db 781 GATTACACACCATGGAACACCTTGTCTAAGTTTGCATAACGCGCAATTTATTTGTACAA 840
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 841 CGCACGCGCAGAGTTGCCCGCAGCGCCACCCCGTTATTGGATTGATCATGTGCAGCG 900
Qy 301 LeuThrProHisProGlnLysGlnAlaTyrluGlyValThrLeuProThrSerValLeu 320
Db 901 TTGACGCCCCCATCCACCGCAAAACAGCGGTATGTGTGACATTACCCACTTCAGTACTG 960
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 961 TTTATTGCGGACACGATACTTAATCTGGCAATCTCGCGCGGCGNACTGGAGCTCAACTGG 1020
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 1021 ACGCTTCCCGGTGAGCGGATAACACGCGCGAGGTGTGAACTGGTGTGAAACGCTGG 1080
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1081 CGTGGCTTAACGATTAACAGCCAGTGGATTGAGTTTCGCTGGTTCCTCAGACTTTTACAG 1140
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 1141 CAGATGGTGATAAAACGCGCTGCTCATTAATACGCGCGCGGAGGTGNAACTGACC 1200
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1201 CTGCGCAGGATGTGAAGAGCGAAATCGCAGCGCATGTGTTCTGTTGGCAGGTTTACGCA 1260
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
Db 1261 ATCGTGAATGAAGCACCGATACCGCGGTGCAAGTTTGAGATCTCATCACCATCACCATCAC 1320

RESULT 3

AAD25460
ID AAD25460 standard; DNA; 1323 BP.

XX AAD25460;

DT 26-MAR-2002 (first entry)

XX Escherichia coli B phytase DNA.

XX Bacterial phytase; K12 appA phytase; protease stability; anabolic;
KW gastrointestinal; nutritional value; feed treatment process; therapy;
KW thermal tolerance; growth performance; alcoholic drink; biopulping;
KW non-alcoholic drink; biobleaching; B phytase; ds.

OS Escherichia coli.

XX Key Location/Qualifiers

FF CDS 1..1323

FT /*tag= a

FT /product= "E. coli B phytase protein"

XX MO200190333-A2.

XX 29-NOV-2001.

Phytase; ds; gene; phytate; appA gene; animal feed; inorganic phosphate; digestion enhancement; transgenic; thermal tolerance; protease stability.

```
Escherichia coli B.
Key      Location/Qualifiers
CDS      1..1323
          /tag= a
          /product= "phytase"
```

US2002136754-A1.

26-SEP-2002.

24-MAY-2001; 2001US-00866379.

13-AUG-1997; 97US-00910798.

13-APR-1999; 99US-00291931.

25-MAY-2000; 2000US-00580515.

(SHOR/) SHORT J M.

(GRAY//) GRAY K A.

(GARR/) GARRETT J

(MATH//) MATHUR E J

Short JM, Kretz K

11

P-PSDB; ADA19446.

Isolated Escherich

nutritional value

1. **Geographical Area** Consider the following:

.....

encoding a phytase enzyme appearing as ADA19450 and having

Also included the *E. coli* appA gene ADA19449 (or an oli-

host cells, a method of improving nutritional value of a phytate-containing feedstuff by contacting the phytate-containing feedstuff with a pure phytase enzyme mutated phytase (where the phytase enzyme catalyzes the liberation of inorganic phosphate from the phytate in the phytate-containing feedstuff), a method to produce an animal feed containing a microbial phytase (comprising culturing the plant cell, plant part or plant under conditions where the nucleotide sequence is expressed and converting the plant cells, plants or plant into a composition for animal feed), a feed composition for animals (comprising the plant seeds, plant cells, plant parts or plants in admixture with a phytate-containing feedstuff), a method to treat a human or an animal able to benefit from digestive enhancement by the activity of an exogenous phytase enzymes comprising administering to the human or animal the plant seed, plant cells, plant parts or plants of a transgenic plant which is modified to contain an expression system which expresses a nucleotide sequence encoding a phytase enzyme, a transgenic non-human organism whose genome comprising a heterologous nucleic acid sequence encoding a polypeptide having phytase activity. The phytase enzyme is useful for improving the nutritional value of phytate-containing feedstuff, in the production of animal feed and for enhancing digestion in humans and animals. The invented method improves thermal tolerance and protease stability. It also improves the feeding value of phytate rich ingredients. The present sequence encodes *E. coli* B wild-type phytase.

Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;

Alignment Scores:			
Pred. No.:	3.15e-202	Length:	1323
Score:	2302.00	Matches:	440
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	8	Gaps:	0
US-10-601-319-2 (1-440) x ADA19445 (1-1323)			
Qy	1	MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla	20
Db	1	ATGAAGCGATCTTAATCCCAATTTATCTCTCTGATTCGGTTAACCCCGCAATCTGCA	60
Qy	21	PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly	40
Db	61	TTTCCTCAGAGTGAGCCGGAGCTGAAGCTGGAAGTGGTGGTGTGATGTCTAGTCCTCATGGT	120
Qy	41	ValArgAlaProThrLysElaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro	60
Db	121	GTGGTGCTCCACCAAGGCCACCGCACTGATGCAAGGATGTCAACCCAGCAGCATGCCCA	180
Qy	61	ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu	80
Db	181	ACCTGGCCGGTAAACTGGGTGGCTGCACCGCGNGGTGGTGGAGCTAATCGCCTATCTC	240
Qy	81	GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro	100
Db	241	GGACATTTACCAACGCCACGCTCTGGTAGCCGACGGATTCCTGGCGAAAAAAGGGCTGCCCG	300
Qy	101	GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu	120
Db	301	CAGTCTGGTCAGGTCGGGATTAATTCGTGATGTCACGAGCGTACCCGTAACACAGCGCGAA	360
Qy	121	AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr	140
Db	361	GCCTTCGCCCGCGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG	420
Qy	141	SerSerProAspProLeuPheAenProLeuLysThrGlyValCysGlnLeuAspAenAla	160
Db	421	TCAGTCTCCCATCCGTTATTTAAATCCCTTAAAAACTGGCGTTTGGCAACTGATTAACCGG	480
Qy	161	AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis	180
Db	481	AACGTGACTGACGGCATCCTCAGCAGGGCAGAGGGTCAATTCCTGACTTTACCGGGCAT	540
Qy	181	ArgGlnThrAlaPheArgGluLeuGluArgValLeuAenPheProGlnSerAenLeuCys	200
Db	541	CGGCMAACGGCGTTTCGCGAATCGGAACGGGTGCTTAATTTCCGCAATCAAACTTGTGC	600
Qy	201	LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu	220
Db	601	CTTAAACGTGAGAAACAGACGAAGCTGTTTATTAAACGAGGCCATTACCATCGGAACCTC	660
Qy	221	LysValSerAlaAspAenValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr	240
Db	661	AAGGTGAGCCCGACAAATGTCCTCAATACCGGTGGTAAAGCCTCGCATCAATGCTGAGC	720
Qy	241	GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr	260
Db	721	GAGATATTTCTCTGCAACAAGCACAGGAATGCGGAGCGCGGGTGGGGAAGGATCACC	780
Qy	261	AspSerHisGlnTrpAenThrLeuLeuSerLeuHisAenAlaGlnPheTyrlLeuLeuGln	280
Db	781	GATTTCAACAGTGGAAACACCTTGCTTAAGTTTGCATAACGCGCAATTTTATTTGCTACA	840
Qy	281	ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla	300
Db	841	CGCACGCCAGAGGTTGCCCGCAGCCGCGCACCCCGTTATTGGATTTGATCATGGCAGCG	900
Qy	301	LeuThrProHisProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu	320

Db 901 TTGAGCGCCCATCCACCGCAAAACAGCGGTATGGTGTGACATTACCCACTTCAGTACTG 960
 Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 Db 961 TTTATTTCGGGACACGATACATAATCTGCCTGCAATCTCGCGCGCGCACTGGAGCTCAACTGG 1020
 Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluAtrgTrp 360
 Db 1021 ACGTTCCCGGTACGCGGATAACACGCCGCCAGGTGGTGAACCTGGTTTGAACGCTGG 1080
 Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 Db 1081 CGTCGGCTAAGGATACAGCCATGGATTGAGTTTCGGTGGTCTTCAGACTTTACAG 1140
 Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 Db 1141 CAGATGGCTGATAAAACGCCGCTGCTATTAATACGCCGCCGCGAGAGGTGAACCTGACC 1200
 Qy 401 LeuAlaGlyCysGluGluAtrgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 Db 1201 CTGGCAGGATGTGAAGACCGAAATGCGCAGGGCATGTGTTGCTGGCAGGTTTACGCA 1260
 Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
 Db 1261 ATCGTGAATGAAGACGCATACCGCGGTGCACTTTGAGATCTCATCACCATCACCATCAC 1320

RESULT 5

ADO50291
 ID ADO50291 standard; DNA; 1323 BP.

AC ADO50291;

XX 29-JUL-2004 (first entry)

DT 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

XX 29-JUL-2004 (first entry)

DR WPI; 2004-374952/35.
 DR P-PSDB; ADO50292.

XX Producing phytase, involves providing nucleic acid derived from bacteria
 PT encoding polypeptide having phytase activity, and expressing nucleic acid
 PT in yeast.

XX Claim 3; SEQ ID NO 1; 74pp; English.

XX The invention relates to a method for producing phytase that involves
 CC providing a nucleic acid encoding phytase derived from a bacteria, and
 CC expressing the nucleic acid in a yeast under conditions that allow
 CC expression of the enzyme in the yeast. The invention also relates to
 CC modified phytase enzyme which has improved thermal tolerance and protease
 CC stability at low pH. The phytase enzyme can be used in foodstuffs to
 CC improve the feeding value of phytate rich ingredients, and in diet of
 CC numerous animals including mammals, fowls and fishes, commercially
 CC significant mammals such as pigs, goats, laboratory rodents, commercially
 CC significant avian species such as chicken, ducks, doves, parrot, etc.,
 CC commercially farmed fish such as guppy, zebrafish, molly, swordtail,
 CC etc., in dough making and baking, in dietary aids for animals. The method
 CC provides easy manufacture of the active ingredient loaded biocompatible
 CC composition, higher yields and loading efficiency. The phytase
 CC incorporated in to the dietary aid is safe for animals. The present
 CC sequence is Escherichia coli B phytase DNA.

SQ Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;

Alignment Scores:
 Pred. No.: 3,15e-202 Length: 1323
 Score: 2302.00 Matches: 440
 Percent Similarity: 100.0% Conservatives: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 12 Gaps: 0

US-10-601-319-2 (1-440) x ADO50291 (1-1323)

Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 Db 1 ATGAAAGCGATCTTAATCCCATTTTATCTCTTCTGATTCCTTACCCGCAATCTGCA 60
 Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
 Db 61 TTCGCTCAGAGTCAGCGGAGCTGAAGCTGAAGTGTGGTGTGATGTGTCGTCATGGT 120
 Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 Db 121 GTGGTGTCTCAACCAAGGCCACGCACTGATGAGGATGTCAACCCAGCGCATGGCCA 180
 Qy 61 ThrTrpProValLysIleuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrlleu 80
 Db 181 ACCTGCGCGTAAACACTGGGTTGGCTGCACACCGCGNGGTGTGAGCTAATCGCTATCTC 240
 Qy 81 GlyHisTyGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 Db 241 GGACATTACCACGCCAGCGCTCTGGTAGCCGCGATGTCTGGCAAAAAGGGCTGCCCG 300
 Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 Db 301 CAGTCTGGTCAGGTTCGGATTATTGCTGATGTCGACGAGCGTACCCGTAACACAGCGCAA 360
 Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 Db 361 GCCTTCGCGCGCGGCTGGCACCCTGCTGTGCAATAACCGTACATACCCAGGAGATACG 420
 Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 Db 421 TCCAGTCCCCGATCCGTTATTATTCCTCTAATAACCTGGCGCTTGGCCAACTGGATAACGCG 480
 Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
 Db 481 AACGTGACTGACGCGATCCTCAGCAGGCGAGGAGGTCAATTGCTGACTTTACCGGCAT 540

Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
 Mather EJ;

QY 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPhleProGlnSerAsnLeuCys 200
 DB 541 CGGCAACCGCGTTTCGCGAAGCTGGAACGGGTGCTTAATTTCCGCAATCAACTTGTGC 600
 QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 DB 601 CTTAAACGTGAGAAACAGACGAAAGCTGTTCTATTAACGACGAGCATTTACCATCGGAATC 660
 QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 DB 661 AAGGTGAGCGCGACGACGATGCTCATTAACCGGTGCGTAAGCTCGCATCAATGCTGACG 720
 QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyValGlyIleThr 260
 DB 721 GAGATATTCTCTCTCAACAGCACAGGAGATGCCGAGCGCGGTGGGAGGATCACC 780
 QY 261 AspSerHisGlnThrAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
 DB 781 GATTCACACCAAGTGGACACCTTGTCTAAGTTTGCATAACGCGCAATTTATTGTGTACAA 840
 QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 DB 841 CGCAGCCAGAGTTGCCGCGAGCGCGCCACCCGTTATTGGATTGATCATGCGAGCG 900
 QY 301 LeuThrProHisProProGlnLysGlnAlaTrpValThrLeuProThrSerValLeu 320
 DB 901 TTGACGCCCCATCCACCGCAAAACAGCGGTATGTGTGACATTACCCACTTCAGTACTG 960
 QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaLeuLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 DB 961 TTTATTGCGGACAGCATACTAATCTGCAAAATCTCGCGCGGCACTGGAGCTCAACTGG 1020
 QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 DB 1021 ACGCTTCGCGGTGACCGGATTAACAGCGCGCGAGTGTGTGAAGTGTGTGAACGCTGG 1080
 QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 DB 1081 CGTCGGCTAAGCGATAACAGCCAGTGGATTGAGTTTGGTGTCTTCAGACTTTACAG 1140
 QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 DB 1141 CAGATGCGTGATAAAACGCGCTGTCAATTAATACGCGCGGAGAGGTGAACCTGACC 1200
 QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 DB 1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGCGCATGTGTGTTGCGCAGGTTTACGCAA 1260
 QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
 DB 1261 ATCGTGATGAAGCAGCATACCGCGGTGAGTTTGAGATCTCATCATCATCATCATCAT 1320

RESULT 6

AEE75412

ID AEE75412 standard; DNA; 1323 BP.

XX AC AEE75412;

XX DT 23-FEB-2006 (first entry)

XX DE E. coli B phytase coding sequence SEQ ID NO: 1.

XX KW ds; coding sequence; thermostable; phytase; protein stabilization;

XX KW pharmaceutical; fertilizer; cramp; cell culture; osteoporosis;

XX KW osteopathic; food.

XX OS Escherichia coli B.

XX FH Key Location/Qualifiers

XX FT 1..1323

XX FT /*tag= a

XX FT /product= "Phytase"

FT unsure 214..216
 FT /*tag= b
 FT /note= "ecodes Arg"

PN US2005281792-A1.

XX PD 22-DEC-2005.

XX PF 01-SEP-2004; 2004US-00933115.

XX PR 13-AUG-1997; 97US-00910798.

XX PR 01-MAR-1999; 99US-00259214.

XX PR 13-APR-1999; 99US-00291931.

XX PR 25-MAY-1999; 99US-00318528.

XX PR 25-MAY-2000; 2000US-00580515.

XX PR 24-MAY-2001; 2001US-00866379.

XX (SHOR/) SHORT J M.

PA (KRET/) KRETZ K A.

PA (GRAY/) GRAY K A.

PA (BART/) BARTON N R.

PA (GARR/) GARRETT J B.

PA (ODON/) O'DONOGHUE E.

PA (BAUM/) BAUM W.

PA (ROBE/) ROBERTSON D E.

PA (ZORN/) ZORNER P.

XX Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, Odonoghue E;

PI Baum W, Robertson DE, Zorner P;

XX WPI; 2006-055022/06.

DR P-PSDB; ABE75413.

XX Formulation useful as dietary supplement for treating, preventing or

PT reversing osteoporosis and bone loss, and preventing muscle cramps,

PT comprises one or more polypeptides having phytase activity.

XX Claim 1; SEQ ID NO 1; 82pp; English.

XX The present sequence is that of a polynucleotide encoding the Escherichia

coli B phytase. The present invention relates to a novel formulation,

CC comprising at least one polypeptide having phytase activity, where the

CC polypeptide is a modified sequence (ABE75412) derived from the

CC Escherichia coli K-12 appA phytase (ABE75419). The modification of the

CC enzyme, comprising a series of mutations, improves the thermal tolerance

CC and protease stability of the protein. The specification also claims a

CC pharmaceutical composition; a kit; an immobilized phytase; a fertilizer

CC or soil additive; a liquid supplement for preventing muscle cramps; a

CC hydrating agent; a tissue culture or cell culture media; and a plant food

CC additive, all comprising the novel phytase, and a method of reducing

CC pollution and increasing nutrient availability in an environment or

CC environmental sample by degrading environmental phytic acid. The novel

CC enzyme is a dietary supplement useful for treating, preventing or

CC reversing osteoporosis or bone loss, and preventing muscle cramps. The

CC liquid supplement is useful for preventing muscle cramps. The method is

CC useful for reducing pollution and increasing nutrient availability in an

CC environment or environmental sample by degrading environmental phytic

CC acid, where the environment or environmental sample comprises a soil or a

CC body of water. The immobilized phytase is useful in foodstuffs for

CC improving the feeding value of phytate rich ingredients.

XX SQ Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 0 U; 1 Other;

Alignment Scores:
 Pred. No.: 3.15e-202 Length: 1323
 Score: 2302.00 Matches: 440
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 15 Gaps: 0

US-10-601-319-2 (1-440) x ABE75412 (1-1323)

Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAAGCGATCTTAATCCCAATTTTATCTCTTCTGATTCGGTTAAACCCCGCAATCTGCA 60
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db 61 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGGTGAATGTTCAGTCGTCTATGGT 120
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 121 GTGCGTGTCTCAACCAAGGCCACGCAACTGATGCAGGATGTCAACCCAGACGCGATGGCCA 180
Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db 181 ACCTGGCCGGTAAACATGGGTGTGCTGACACCGGNGGTGTGAGCTAATCGCCTATCTC 240
Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 241 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGGAATGTCTGGCGAAAAGGGCTGCCCCG 300
Qy 101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 301 CAGTCTGCTCAGGTCCGGATTAATGCTGATGTCGACGAGCGTACCCGTAAACACAGGCGAA 360
Qy 121 AlaPheAlaIleGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCCTTCGCGCGCGGCTGGCACCTGTGCTGCAATAACCGTACATACCCAGCGAGATACG 420
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCCAGTCCCGATCCGTTATTAATCTTAATAAAGCTGGCGTTGCCCACTGGATAACGCG 480
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGACTACGCGATCTCTCAGCAGCGCAGGAGGGTCAATTGCTGACTTTTACCGGGCAT 540
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 541 CGGCAAAACGGGGTTTCGGAACCTGGAAACGGGTGTGTAATTTTCGCAATCAAACTTGTGC 600
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAAACGTGGAACAACAGGACGAAGCTGTTCATTACGCGAGGATTAACCATCGGAATC 660
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGAGCGCGCAATGTCTATTAAACCGGTGCGGTAAAGCCTCGCATCAATGCTGACG 720
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Db 721 GAGATATTCTCTGCAACAAAGCAACAGGAATCCGAGCGGGGTGGGAAGGATCACC 780
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 781 GATTCACACCACTGGAACACCTTCCTAAGTTTGATTAACCGCAATTTATTTGCTACAA 840
Qy 281 ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 841 CGCACGCCAGAGGTGCGCGAGCGCGCCACCCCGCTTATTGGATTTTGATCATGCGACG 900
Qy 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 901 TTGACGCCCATTCACCGCAAAAACAGCGGTATGGTGTGACATTACCACCTTCAGTACTG 960
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 961 TTTATTCCGCGACACGATACTAATCTGCAAAATCTCGCGCGCGCACTGGAGCTCAACTGG 1020
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 1021 ACGCTTCGCGTCAGCGGATTAACACCGCGCAGGTGTGAACTGGTGTGTTGAACGCTGG 1080
Qy 361 ArgArgLeuSerAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380

Db 1081 CGTCGGCTAAGCGATACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1140
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 1141 CAGATCGGTGATAAAACGCCGCTGTCAATTAATAACGCCGCCGCGAGAGGTGAAACTGACC 1200
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1201 CTGCGCAGGATGTGAAGACGCAAAATGCGCAGGCAATGTGTTCGTTGGCAGGTTTTTACGCA 1260
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
Db 1261 ATCGTGAATGAAGCACGCATACCGCGTGCAGTTTGAGATCTCATCACCATCACCATCAC 1320
RESULT 7
AD050303
ID AD050303 standard; DNA; 1323 BP.
XX
AC AD050303;
XX
DT 29-JUL-2004 (first entry)
XX
DE Escherichia coli W phytase 875PH2 mutant DNA.
XX
KW Phytase; bacteria; thermal tolerance; protease stability; foodstuff;
KW phytate; animal feed; fish feed; dough; baking; mutant; gene; ds.
XX
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT CDS 1..1323
FT FT /*tag= a
FT FT /product= "Mutant phytase enzyme"
FT FT 214..216
FT FT /*tag= b
FT FT /note= "Encodes Arg"
FT FT replace(478,G)
FT FT /*tag= c
FT FT replace(527,A)
FT FT /*tag= d
FT FT replace(893,T)
FT FT /*tag= e
FT FT replace(895,G)
FT FT /*tag= f
XX
PN US2004091968-A1.
XX
PD 13-MAY-2004.
XX
PF 20-JUN-2003; 2003US-00601319.
XX
PR 13-AUG-1997; 97US-00910798.
PR 01-MAR-1999; 99US-00259214.
PR 13-APR-1999; 99US-00291931.
PR 25-MAY-1999; 99US-00318528.
PR 25-MAY-2000; 2000US-00580515.
PR 24-MAY-2001; 2001US-00866379.
XX
PA (SHOR/) SHORT J M.
PA (KRET/) KRETZ K.
PA (GRAY/) GRAY K A.
PA (BART/) BARTON N R.
PA (GARR/) GARRETT J B.
PA (ODON/) O' DONOGHUE E.
PA (MATH/) MATHER E J.
XX
PI Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI Mather EU;
XX
XX WPI; 2004-374952/35.
DR P-PSDB; AD050292, AD050304.
XX

PT Producing phytase, involves providing nucleic acid derived from bacteria
 PT encoding polypeptide having phytase activity, and expressing nucleic acid
 XX in yeast.

PS Disclosure; Page; 74pp; English.

XX The invention relates to a method for producing phytase that involves
 CC providing a nucleic acid encoding phytase derived from a bacteria, and
 CC expressing the nucleic acid in a yeast under conditions that allow
 CC expression of the enzyme in the yeast. The invention also relates to
 CC modified phytase enzyme which has improved thermal tolerance and protease
 CC stability at low pH. The phytase enzyme can be used in foodstuffs to
 CC improve the feeding value of phytate rich ingredients, and in diet of
 CC numerous animals including mammals, fowls and fishes, commercially
 CC significant mammals such as pigs, goats, laboratory rodents, commercially
 CC significant avian species such as chicken, ducks, doves, parrot, etc.,
 CC commercially farmed fish such as guppy, zebrafish, molly, swordtail,
 CC etc., in dough making and baking, in dietary aids for animals. The method
 CC provides easy manufacture of the active ingredient loaded biocompatible
 CC composition, higher yields and loading efficiency. The phytase
 CC incorporated in to the dietary aid is safe for animals. The present
 CC sequence is Escherichia coli W phytase mutant DNA. Note: The present
 CC sequence is not shown in the specification but has been derived from
 CC Escherichia coli B phytase DNA ADO50291.

XX SQ Sequence 1323 BP; 324 A; 353 C; 356 G; 289 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 2 2e-200 Length: 1323
 Score: 2282.00 Matches: 436
 Percent Similarity: 99.3% Conservat: 1
 Best Local Similarity: 99.1% Mismatches: 3
 Query Match: 99.1% Indels: 0
 DB: 12 Gaps: 0

US-10-601-319-2 (1-440) x ADO50303 (1-1323)

Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 Db 1 ATGAAGCGATCTTAATCCCAATTTTATCTCTCTGATTCCGTAAACCCCGCAATCTGCA 60
 Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
 Db 61 TTCGCTCAGAGTCAGCGGAGCTGAAGCTGGAAGATGTGTGATTGTTCAGTCGTCATGGT 120
 Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAapValThrProAspAlaTrpPro 60
 Db 121 GTGGCTGCTCCAAACAGGCCAGCACTGATGCAGGATGTACCCACAGCGCATGGCCA 180
 Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
 Db 181 ACCTGGCCGGTAAACTGGGTGTGCTGACACCGCGNGGTGTGAGCTAATCGCTATCTC 240
 Qy 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 Db 241 GGACATTACCAACGCCAGCGTCTGTAGCCACGAGATTGCTGGCGAAAGGGCTGCGCG 300
 Qy 101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 Db 301 CAGTCTGTGTCAGGTCCGATTAATGTGATGTGACAGAGCGTACCCGTAAACAGGGCGAA 360
 Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 Db 361 GCCTTCGCGCGCGGCTGGCACTGCTGTGCAATACCGTACATACCCAGCGAGATACG 420
 Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 Db 421 TCCAGTCCCGATCCGTTATTTAATCTCTAAATAAATGGCGTTTGCCTAACTGATAACTCG 480
 Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 Db 481 AACGTGACTACGCGATCTCTCAGCAGGCGCAGGGGTCAATTTGCTGGCTTTTACGGGCAT 540

Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
 Db 541 CGGCAACCGGCTTTCGCGAATCTGGAACGGGTCTTAATTTCCGCAATCAAACTTGTGC 600
 Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 Db 601 CTTAAACGTGAGAAACAGGACGAAGCTGTTCATTAAACGCGGCAATTACCATCGAACTC 660
 Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 Db 661 AAGGTGAGCGCGCAATGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGAG 720
 Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
 Db 721 GAGATATTCTCTGCAACAGCAGAGGAATGCGGAGCCGGGTGGGGAAGGATCACC 780
 Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
 Db 781 GATTCAACACCAAGTGGAAACACCTTCTTAAGTTTGCATAACGGCGCAATTTTATTGCTACA 840
 Qy 281 ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 Db 841 CGCAGCGCAGAGGTGCCCGCAGCGCCGCCACCCCGTTATTGGATTGATCAAGACAGCG 900
 Qy 301 LeuThrProHisProProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
 Db 901 TTGACGCCCATCCACCGCAAAAACAGGCGTATGTGTGACATTACCCCACTTCAGTACTG 960
 Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 Db 961 TTTATTCCGCGACACGATACCTTAATCTGGCAATCTCGCGCGGCACTGGAGCTCAACTGG 1020
 Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 Db 1021 ACGTTTCCGCTCAGCGGATTAACAGCGCCGAGTGTGTGATGTGTGTGAAACGCTGG 1080
 Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 Db 1081 CGTCGGCTAAGCGATAACAGCCAGTGGATTTCAGGTTTCGCTGGTCTTCCAGACTTTTACAG 1140
 Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 Db 1141 CAGATGGGTGATAAAGCCCGCTGTCTATTAAATACGCCGCCCGGAGAGGTGAAACTGACC 1200
 Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 Db 1201 CTGGCAGATGTGAAGACGGAATGCGCAGGCGATGTTCGTTGGCAGGTTTACGCAA 1260
 Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
 Db 1261 ATCGTGAATGAAGCAGCATACCGCGTGCAGTTTGAGATCTCATCACCATCACCATCAC 1320
 RESULT 8
 ADO50301 standard; DNA; 1323 BP.
 ID ADO50301
 XX ADO50301;
 AC ADO50301;
 DT 29-JUL-2004 (first entry)
 XX Kangaroo rat Escherichia coli phytase 872PH1 mutant DNA.
 DE Phytase; bacteria; thermal tolerance; protease stability; foodstuff;
 KW phytate; animal feed; fish feed; dough; baking; mutant; gene; ds.
 XX Escherichia coli.
 OS Escherichia coli.
 FH Key Location/Qualifiers
 FT CDS 1..1323
 FT /*tag= a
 FT /product= "Phytase mutant enzyme"
 FT 214..216
 FT /*tag= b

```

FT mutation /note= "Encodes Arg"
FT replace(527,A)
FT /*tag= c
FT mutation replace(893,T)
FT /*tag= d
FT mutation replace(895,G)
FT /*tag= e
FT mutation replace(934,G)
FT /*tag= f
FT
XX US2004091968-A1.
XX
XX 13-MAY-2004.
XX
XX 20-JUN-2003; 2003US-00601319.
XX
XX 13-AUG-1997; 97US-00910798.
XX 01-MAR-1999; 99US-00259214.
XX 13-APR-1999; 99US-00291931.
XX 25-MAY-1999; 99US-00318528.
XX 25-MAY-2000; 2000US-00580515.
XX 24-MAY-2001; 2001US-00866379.
XX
XX (SHOR/) SHORT J M.
XX (KRETZ/) KRETZ K.
XX (GRAY/) GRAY K A.
XX (BART/) BARTON N B.
XX (GARR/) GARRETT J B.
XX (ODON/) O'DONOGHUE E.
XX (MATH/) MATHER E J.
XX
XX Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
XX Mather EJ;
XX
XX WPI: 2004-374952/35.
XX P-PSDB; ADO50302.
XX
XX Producing phytase, involves providing nucleic acid derived from bacteria
XX encoding polypeptide having phytase activity, and expressing nucleic acid
XX in yeast.
XX
XX Disclosure; Page; 74pp; English.
XX
XX The invention relates to a method for producing phytase that involves
XX providing a nucleic acid encoding phytase derived from a bacteria, and
XX expressing the nucleic acid in a yeast under conditions that allow
XX expression of the enzyme in the yeast. The invention also relates to
XX modified phytase enzyme which has improved thermal tolerance and protease
XX stability at low pH. The phytase enzyme can be used in foodstuffs to
XX improve the feeding value of phytate rich ingredients, and in diet of
XX numerous animals including mammals, fowls and fishes, commercially
XX significant mammals such as pigs, goats, laboratory rodents, commercially
XX significant avian species such as chicken, ducks, doves, parrot, etc.,
XX commercially farmed fish such as guppy, zebrafish, molly, swordtail,
XX etc., in dough making and baking, in dietary aids for animals. The method
XX provides easy manufacture of the active ingredient loaded biocompatible
XX composition, higher yields and loading efficiency. The phytase
XX incorporated in to the dietary aid is safe for animals. The present
XX sequence is Kangaroo rat Escherichia coli phytase mutant DNA. Note: The
XX present sequence is not shown in the specification but has been derived
XX from Escherichia coli B phytase DNA ADO50291.
XX
XX Sequence 1323 BP; 325 A; 353 C; 356 G; 288 T; 0 U; 1 Other;
SQ

```

Alignment Scores:

Pred. No.:	4.16e-200	Length:	1323
Score:	2279.00	Matches:	436
Percent Similarity:	99.1%	Conservative:	0
Best Local Similarity:	99.1%	Mismatches:	4
Query Match:	99.0%	Indels:	0
DB:	12	Gaps:	0

US-10-601-319-2 (1-440) x ADO50301 (1-1323)

1 MetLysAlaIleLeuLeuProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
 1 ATGAAAGCGATCTTAATCCCATTTTATCTCTCTGATTCCGTTAAACCCGCAATCTGCA 60
 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
 61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAGTGTGGTGAATGTTCATCGTCATGGT 120
 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 121 GTGCGTGTCTCAACCAAGCCAGCCAGCACTGATCAGGATGTCAACCCAGAGCATGGCCA 180
 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
 181 ACCTGGCCGGTAAACCTGGGTTGGCTGACACCCGCGNGGTGGTGAGCTTAATCGCTATCTC 240
 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
 241 GGACATTTACCACGCCAGCGCTCTGGTAGCCGCGGATTTGCTGGCGAAAAAGGGCTGCCCG 300
 101 GlnSerGlyGlnValAlaIleLeuAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 301 CAGTCTGGTCAGGTTCGGATTATTGCTGATGTCGACGAGCGTACCCGTTAAACAGGCGAA 360
 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 361 GCCTTCGCCCGCGGCTGGCACCTGACTGTGCAATAACCGTATACATACCCAGGAGATACG 420
 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 421 TCAGTCCCGGATCCGTTATTAAATCTCTAAAAACGTGGCGTTTGCCAACTGGATACGCG 480
 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 481 AACGTGACTGACGCGATCTCTACAGGCGCAGGAGGGTCAATTGCTGCTGCTTTTACCGGSCAT 540
 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
 541 CGGCAAAACGGCGTTTCGCGAACTGGAAACGGGTCTTAATTTTCGCAATCAAACTTTGTGC 600
 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTACGCGGCACTTACCATTCCGAATC 660
 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 661 AAGGTGAGCGCCGCAATGCTCATTAACCGGTGCGGTAAAGCTCGCATCAATGCTGACG 720
 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
 721 GAGATATTTCTCTGCAACAGCACAGGGAATGCGGAGCCGGGTGGGGAGGATCACCC 780
 261 AspSerHisGlnTrpAsnThrLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
 781 GATTACACCAAGTGGAAACCTTGTCAAGTTTGCATTAACCGCAATTTATTTCTCTACAA 840
 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 841 CGCACGCCAGAGGTTCGCCGCGCAGCCGCCACCCCGTTATTGGATTGATCAAGACAGCG 900
 301 LeuThrProHisProProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
 901 TTGACGCCCCCATCCACGCAAAAACAGGCGGTATAGTGTGATCATTTACCCATCTCAGTACTG 960
 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 961 TTTATTGCCGGACACGATATACTATCTGGCAAAATCTCGCGCGCGCAGCTGGAGCTCAACTGG 1020
 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 1021 ACGCTTCCCGGTGAGCCGGATACACGCCGCCAGGTGGTGAACACTGGTGTGTTTGAACGCTGG 1080

QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 Db 1081 CGTCGGCTAAAGCGATAACAGCCAGTGGATTGAGGTTTCGCTGCTCCAGACTTTACAG 1140
 QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 Db 1141 CAGATGCGTGATAAACCGCGCTGTCATTAAATACGCCCGCGAGAGGTGAACTGACC 1200
 QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 Db 1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGCGCATGTGTCGTTGGCAGGTTTACGCAA 1260
 QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
 Db 1261 ATCGTGAATGAAGCAGCATACCGCGCTGCAGTTTGAGATCTCATCACCATCACCATCAC 1320

RESULT 9

ABK12514

ID ABK12514 standard; DNA; 1299 BP.

XX AC

ABK12514;

XX XX

05-JUN-2002 (first entry)

XX DT

DE DNA encoding phytase associated protein.

XX XX

Phytase; gene; ds.

XX XX

Unidentified.

XX XX

Key Location/Qualifiers

FH CDS 1..1299

FT FT /*tag= a

FT FT /product= "Phytase associated protein"

XX XX

KR99086028-A.

XX XX

15-DEC-1999.

XX XX

25-MAY-1998; 98KR-00018810.

XX XX

25-MAY-1998; 98KR-00018810.

XX XX

(WO0J-) WO0JIN CO LTD.

XX XX

Bae HD, Forceburgh CW, Goloben S, Cheng KJ;

XX XX

WPI; 2000-645078/62.

DR DR

P-PSDB; AAU77775.

XX XX

Novel phytase gene, recombinant phytase and usage thereof.

XX XX

Claim 1; Fig 2; 10pp; Korean.

XX PS

The invention relates to a novel phytase gene, a recombinant phytase gene

CC CC

and their uses. This sequence encodes a phytase associated protein,

CC CC

described in the invention

XX XX

Sequence 1299 BP; 317 A; 345 C; 355 G; 282 T; 0 U; 0 Other;

SQ

Alignment Scores:

Pred. No.: 4.63e-196 Length: 1299

Score: 2235.00 Matches: 430

Percent Similarity: 99.5% Conservative: 0

Best Local Similarity: 99.5% Mismatches: 2

Query Match: 97.1% Indels: 0

DB: 3 Gaps: 0

US-10-601-319-2 (1-440) x ABK12514 (1-1299)

QY

1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20

Db

1 ATGAAGCGATCTTAATCCCAATTTTATCTCTTCTGATTCCGTTTAAACCCGCAATCTGCA 60

QY 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
 Db 61 TTCGCTCAGAGTGGAGCGAGCTGAAGCTGGAAGTGTGGTATTGTTCAGTCGTATGGT 120

QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 Db 121 GTGGGTGCTCCAACAGGCCACGCACTGATGCGAGGATGTCAACCCAGAGCGCATGGCCA 180

QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGlyLeuIleAlaTrpLeu 80
 Db 181 ACCTGGCCGGTAAACCTGGGTGGCTGACACCGCGCGGTGGTGAAGTAAATCGCCCTATCTC 240

QY 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
 Db 241 GGACATTACCAACGCCAGCGTCTGGTAGCCGCGGATTGCTGGCGAANAAGGCTGCCCG 300

QY 101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 Db 301 CAGTCTGGTCAAGTCCGCAATTATTGCTGATGTCAGCAGCGTACCCGTAANAACAGCGCAA 360

QY 121 AlaPheAlaIleGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 Db 361 GCCTTCGCGCGCGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGCGCATAGC 420

QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 Db 421 TCCAGTCCCGATCCGTTATTATTAACTCTAANAACCTGGCGTTTGCCAATCGATNAACGG 480

QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 Db 481 AACGTGACTGACGCGATCCTCAGCAGGCGCAGGAGGTCAATTTGCTGACTTTACCGGCAT 540

QY 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
 Db 541 CGGCAAAACGGCGTTTCGCAACTGGAACGGGTGCTTAATTTTCGCAATAAACAATTTGTC 600

QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 Db 601 CTTAAACGTGGAANAACAGGACGGAAGCTGTTCAATTAACGCGAGGCATTAACCATCGGAATC 660

QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 Db 661 AAGGTGAGCGCGCAATGTCTCAATTAACCGGTGCGTGAAGCCTCGCATCAATGCTGACG 720

QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
 Db 721 GAGATATTTCTCTGCAACAGCAAGGAAATGCCGAGCGCGGGTGGGAAGGATCACC 780

QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
 Db 781 GATTCAACCCAGTGAACACCTTGTAAAGTTTGCATAACGCGCAATTTTATTGCTACAA 840

QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 Db 841 CGCACGCGCAGAGGTGCGCGCAGCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG 900

QY 301 LeuThrProHisProProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
 Db 901 TTGACGCCCATCCACCGCAAAAACAGCGGTATGGTGTGACATTACCCACCTTCAGTGTCTG 960

QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 Db 961 TTTATCGCGCGACACGATACTTAATCTGCAAAATCTCGCGCGCGCATCTGGAGCTCAACTGG 1020

QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 Db 1021 ACGCTTCCGGTCAGCCGATAACACGCCCGCAGGTGGTGAACCTGGTGTGTTGAACGCTGG 1080

QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 Db 1081 CGTCGGCTAAGCGATAAACGCCAGTGGATTTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1140

QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 Db 1141 CAGATGCGTATAAAACGCCGCTGCTCAATTAATACGCCGCCGAGAGGTGAACCTGACC 1200
 QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 Db 1201 CTGGCAGGATGTGAAGAGCGAATGCGCAGGCAATGTTCTGTTGGCAGGTTTACGCAA 1260
 QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
 Db 1261 ATCGTGAATGAAGCAGCATACCGCGTGCAGTTTG 1296

RESULT 10

ACAL9297
 ID ACA19297 standard; DNA; 1299 BP.

XX
 AC ACA19297;

XX
 DT 19-JUN-2003 (first entry)

XX
 DE Prokaryotic essential gene #954.

XX
 KW Antisense; ds; prokaryotic essential gene; cell proliferation;
 KW drug design; gene.

XX
 OS Escherichia coli.

XX
 PN WO200277183-A2.

XX
 PD 03-OCT-2002.

XX
 PF 21-MAR-2002; 2002WO-US009107.

XX
 PR 21-MAR-2001; 2001US-00815242.

PR
 06-SEP-2001; 2001US-00948993.

PR
 25-OCT-2001; 2001US-0342923P.

PR
 08-FEB-2002; 2002US-00072851.

PR
 06-MAR-2002; 2002US-0362695P.

XX
 PA (ELIT-) ELITRA PHARM INC.

XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX
 DR WPI; 2003-029926/02.

DR
 P-PSDB; ABUI5427.

XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.

XX
 PS Claim 14; SEQ ID NO 7167; 1766pp; English.

XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of

CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The present sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX
 SQ Sequence 1299 BP; 317 A; 345 C; 355 G; 282 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.63e-196 Length: 1299
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.5% Conservative: 0
 Best Local Similarity: 99.5% Mismatches: 2
 Query Match: 97.1% Indels: 0
 DB: 8 Gaps: 0

US-10-601-319-2 (1-440) x ACA19297 (1-1299)

QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 Db 1 ATGAAAGCGGATCTTAATCCCATTTTATCTCTTCTGATTCCTGTTAAACCCGCAATCGCA 60
 QY 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
 Db 61 TTCGTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGGTATTGTTCAGTCGTCAATG 120
 QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 Db 121 GTGCGTGCTCCAAACCAAGGCCAGCAACTGATCAGGATGTCACCCAGACGCGATGCCA 180
 QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyLeu 80
 Db 181 ACCTGGCCGGTAAACCTGGGTTGGCTGACACCCGCGGTGGTGAATTCGCTATCTC 240
 QY 81 GlyHisTyrglnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 Db 241 GGACATTACCAACGCCAGCGCTCTGGTAGCCGACGGATTGTCGGGAAAGGGCTGCCG 300
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 Db 301 CAGTCTGGTCAGGTCGCGATTATTGCTGATGTGACGAGCGTACCCGCTAAACACAGGCGAA 360
 QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 Db 361 GCCTTCGCCCGCGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCGAGATACG 420
 QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 Db 421 TCCAGTCCCGATCGTTATTTATCTCTAAACCTGGCGTTTGGCAACTGGATAACGCG 480
 QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 Db 481 AACGTGACTCAGCGGATCTCTCAGCAGGCGAGGAGGTCAATTTGCTGACTTTTACCGGCA 540
 QY 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
 Db 541 CGGCAACCGCGGTTTCGCAACTGGAAACGGGTGCTTAATTTTCGCAATCAAACTTTGTC 600
 QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 Db 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCTTAACGCGAGGCAATTACCATCGGAATC 660
 QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 Db 661 AAGGTGAGCGCGCACAAATGCTCATTAACCGCGTGGTGAAGCCCTCGCATCAATCTGACG 720
 QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260

Db 721 GAGATATTCTCTCTCAACAGCAGAGGAATGCCGGAGCCGGGTGGGAAGGATCACC 780
 Qy 261 AppSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrluLeuGln 280
 Db 781 GATTTCACACCACTGGAACACCTTGTAGTTTGCATAACGGCAATTTATTTGCTACAA 840
 Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 Db 841 CGCACGCCAGAGTTGCCCGCAGCCGCCACCCCGTTATTAGATTTCATCAAGACACGG 900
 Qy 301 LeuThrProHisProGlnLysGlnAlaTyrlGlyValThrLeuProThrSerValLeu 320
 Db 901 TTGACGCCCATCCACCCCAAAACAGGCGTATGTTGATACATTACCCCACTTCAGTGTG 960
 Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 Db 961 TTTATCGCGGACACGACGATCTAATCTGGCAATCTCGCGCGGCACTGGAGCTCAACTGG 1020
 Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 Db 1021 ACGCTTCGCGTCAGCCGGGATAACAGCCGCCAGGTGTGAATGGTGTGTTGAACGCTGG 1080
 Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 Db 1081 CGTCGGCTTAAGCGATAACAGCCAGTGGATTTCAGGTTTCGCTGGTCTCCAGACTTTACAG 1140
 Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 Db 1141 CAGATGCGTGATTAACCCCGCTGTTCATTAATACGCCCGCGAGAGGTGAATGACC 1200
 Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 Db 1201 CTGGCAGGATGTGAAGACGGAATCGCGAGGCGCATGTGTTGTTGGCAGGTTTACGCAA 1260
 Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
 Db 1261 ATCTGGAATGAACGACGATACCGCGCGTGCAGTTTG 1296

RESULT 11

ID ADL16137 standard; DNA; 1299 BP.

XX AC ADL16137;

XX DT 06-MAY-2004 (first entry)

XX DE Escherichia coli K-12 AppA phytase gene.

XX KW AppA phytase; modified phytase; K-12; animal feed; feed additive;
 KW phosphorus; phytate; myo-inositol hexakisphosphate; food production;
 KW wet grain milling; dental care product; bone resorption; osteoporosis;
 KW osteopathic; kidney stone; metal removal; waste treatment;
 KW cleaning product; rust removal product; gene; ds.

XX OS Escherichia coli; K-12.

XX PH Key Location/Qualifiers

FT CDS 1..1299
 FT /*tag= a
 FT /product= "AppA phytase"

XX WO2004015084-A2.

XX PD 19-FEB-2004.

XX XX 11-AUG-2003; 2003WO-US025058.

XX XX 12-AUG-2002; 2002US-0403330P.

XX PA (GENW) GENENCOR INT INC.

XX PI Blattmann BO, Darzins A, Davis JM, Encell LP, Morrison TB;

PI Mrachko GT, Schellenberger V;

XX WPI: 2004-169498/16.

DR P-PSDB; ADL16129.

XX Novel recombinant phytase having modified phytase activity comprising
 PT modification of amino acid residues in mature Escherichia coli phytase
 PT designated EBC1082, useful as feed additive.

XX Claim 19; Fig 21; 107pp; English.

XX The invention relates to recombinant modified AppA phytases having
 CC modified phytase activity. The AppA phytases are derived from a mature
 CC Escherichia coli AppA phytase designated EBC1082 (see ADL16120), and
 CC comprise substitutions at least one of residues 26, 43, 46, 54, 73, 113,
 CC 125, 184, 228, 384 or 410, or at a residue located within 5 residues of
 CC those listed above. The invention also encompasses polynucleotides
 CC encoding a modified AppA phytase and which further encodes a secretion
 CC signal sequence operable in Bacillus sp.; expression constructs, vectors
 CC and host cells (preferably Bacillus subtilis) comprising the
 CC polynucleotide sequences; and a method for the production of the modified
 CC AppA phytases. The invention also discloses a method of producing an
 CC enzyme having phytase activity; a method of producing a heterologous
 CC polypeptide having phytase activity in Bacillus subtilis; reducing the
 CC level of phosphorus in animal manure; a composition comprising a modified
 CC AppA phytase; and animal feed comprising cells, spores or plant parts,
 CC including seeds, capable of expressing a modified AppA phytase. The
 CC modified AppA phytases are useful as an animal feed additive,
 CC particularly in feeds for non-ruminant livestock such as pigs and poultry
 CC which lack the digestive enzymes for extracting phosphorus from phytate
 CC (myo-inositol hexakisphosphate), a substance found in plant seed which is
 CC an important source of phosphorus. The modified AppA phytases also have
 CC many other agricultural, industrial, medical and nutritional
 CC applications. For example, they can be used in the isolation and
 CC recovery of rare metals to produce lower phosphate homologs of phytate,
 CC which may be used in toothpaste and other dental care products as well as
 CC potential treatments or preventatives of bone resorption (e.g., in
 CC osteoporosis) and renal calculi (kidney stones); in food production
 CC (e.g., tofu and nutritional supplements); in wet grain milling; in
 CC cleaning, personal care, medical and nutritional products; in rust
 CC removal products; and in the removal of metals and other polycations from
 CC such diverse materials as waste products and carbonated beverages. The
 CC present sequence represents Escherichia coli strain K-12 AppA phytase
 CC gene.

XX SQ Sequence 1299 BP; 317 A; 345 C; 355 G; 282 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.63e-196 Length: 1299
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.5% Conservative: 0
 Best Local Similarity: 99.5% Mismatches: 2
 Query Match: 97.1% Indels: 0
 DB: 12 Gaps: 0

US-10-601-319-2 (1-440) x ADL16137 (1-1299)

Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 Db 1 ATGAAGCGGATCTTAATCCCATTTTATCTCTTCTGATTCGCTTAACCCGCAATCTGCA 60
 Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
 Db 61 TTGCTCAGAGTGAGCGGAGCTGAAGCTGAAGTGTGGTGTATTGTTCAGTCGTCACTGT 120
 Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 Db 121 GTGCGTGTCCAACCAAGGCCACGCAACTGATCAGGATGTCACTCAGCATGATGATGCA 180
 Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrlu 80
 Db 181 ACCTGGCCGGTAAAACTGGGTGTGGCTGACACCCGCGGTGGTGGAGCTAAATCGCTATCTC 240

Qy 81 GlyHisTyrGlnArgGlnArgLeuValalaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db |||||
Qy 241 GGACATTACCAACCCAGCGCTGCTGGTAGCCAGCGATTGCTGGCGAAAAAGGGCTGCCCG 300
Db |||||
Qy 101 GlnSerGlyGlnValAlaIleAlaAspValaAspGluArgThrArgLysThrGlyGlu 120
Db |||||
Qy 301 CAGTCTGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAACACAGCGGAA 360
Db |||||
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db |||||
Qy 361 GCCTTCGCGCGCGCTGGCACCTGACTGTGCAATACCGTACATACCGACGAGATACG 420
Db |||||
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db |||||
Qy 421 TCCAGTCCCGATCCGTTATTATTCCTCAAAACTGCGCTTCCCACTGGATAACGCG 480
Db |||||
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db |||||
Qy 481 AACGTGACTGACGCGATCCTCAGCAGGCGAGGAGGTCAATTGCTGACTTTACCGGCAT 540
Db |||||
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db |||||
Qy 541 CGGCAAAACGGCGTTTCGCAACTGGAACGGGTCTTAATTTTCCGCAATCAAACTTGTC 600
Db |||||
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db |||||
Qy 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAAACGGTGGATACCATCGGAACCTC 660
Db |||||
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db |||||
Qy 661 AAGGTGAGCGCGACATGCTCATTAACCGGTGCGGTAGCGCTCGCATCAATGCTGACG 720
Db |||||
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
Db |||||
Qy 721 GAGATATTCTCTCTGCAACAGCACAGGGAATGCCGAGCGCGGTGGGAAGATCAC 780
Db |||||
Qy 261 AsnSerHisGlnTyrAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db |||||
Qy 781 GAITTCACACAGTGGAAACCTTGCTAAGTTTGCATTAACGCGCAATTTTATTTCGTACAA 840
Db |||||
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db |||||
Qy 841 CGCAGCGCAGAGGTGTCGCCGAGCGCGCCACCCCGTTATTAGATTGATCAAGACAGCG 900
Db |||||
Qy 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db |||||
Qy 901 TTGACGCCCATCCACCGCAAAACAGGCGGTATGCTGACATTACCCACTTCAGTGCTG 960
Db |||||
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTyr 340
Db |||||
Qy 961 TTTATCCCGGACACGATACCTAATCTGGCAATCTCGCGCGGACCTGGAGCTCAACTGG 1020
Db |||||
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTyr 360
Db |||||
Qy 1021 ACGTCTCCGCTCAGCGGATAACACCGCCAGGTGTGTGAACCTGGTGTGAACGCTGG 1080
Db |||||
Qy 361 ArgArgLeuSerAspAsnSerGlnTyrIleGlnValSerLeuValPheGlnThrLeuGln 380
Db |||||
Qy 1081 CGTCGGCTAAGCGGATAACAGCAGTGGATTTCAGGTTTTCGCTGGTCTTCCAGACTTTACAG 1140
Db |||||
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db |||||
Qy 1141 CAGATGCGTATAAAGCGCGCTGTCAATTAATACGCGCGCGGAGGAGGTAACACTGACC 1200
Db |||||
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db |||||
Qy 1201 CTGGCAGGATGTGAAGACGGAATGCGCAGGCGATGTTGCTGTCGTCAGGTTTACGCCAA 1260
Db |||||
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db |||||
Qy 1261 ATCGTGAATGAAGCAGCAGATACCGCGCTGCAGTTTG 1296
Db |||||

RESULT 12

AAD25463
ID AAD25463 standard; DNA; 1901 BP.
XX
AC AAD25463;
DT 26-MAR-2002 (first entry)
XX
DE Escherichia coli appA phytase wild type DNA.
XX
KW Bacterial phytase; K12 appA phytase; protease stability; anabolic;
gastrointestinal; nutritional value; feed treatment process; therapy;
thermal tolerance; growth performance; alcoholic drink; biopulping;
non-alcoholic drink; biobleaching; ds.
XX
OS Escherichia coli.
XX
Key Location/Qualifiers
FT CDS
FT 188..1486
FT /*tag= a
FT /product= "E. coli appA phytase protein"
XX
FN WO200190333-A2.
XX
PD 29-NOV-2001.
XX
PF 24-MAY-2001; 2001WO-US017118.
XX
PR 25-MAY-2000; 2000US-00580515.
XX
PA (DIVE-) DIVERSA CORP.
XX
PI Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, O'donoghue E;
XX
DR WPI; 2002-083108/11.
DR P-PSDB; AAE15807.
XX
PT New bacterial phytase for e.g. improving the nutritional value of phytate
-containing foodstuffs and subsequently improving the growth performance
of an organism that consumes it, or in treating animal digestive systems.
XX
PS Claim 42; Fig 7; 170pp; English.
XX
CC The patent discloses recombinant bacterial phytase from Escherichia coli
K12 appA phytase. The enzyme has phytase activity and improved thermal
tolerance when compared with wild-type phytase. It has improved protease
stability at low pH. The recombinant phytase is useful for improving the
nutritional value of phytate-containing foodstuffs and subsequently
improving the growth performance of an organism that consumes it, in
treating animal digestive systems, in feed treatment processes and for in
vitro purposes related to research, discovery and development. They are
also used for generating recombinant digestive system life forms, for
producing or manufacturing alcoholic and non-alcoholic drinks based on
the use of moulds, grains and/or plants, in biopulping and bio- bleaching
where a reduction in the use of environmentally harmful chemicals that
are traditionally used in the pulp and paper industry is desired and in
the reduction or possible elimination of the need for mineral
supplements, enzymes or therapeutic drugs for animals from the daily feed
thus increasing the amount calories and nutrients present in the feed.
CC The present sequence is a DNA encoding E. coli appA phytase wild type
protein
XX
SQ Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 7.83e-196 Length: 1901
Score: 2235.00 Matches: 430
Percent Similarity: 99.5% Conservative: 0
Best Local Similarity: 99.5% Mismatches: 2
Query Match: 97.1% Indels: 0
DB: 6 Gaps: 0

US-10-601-319-2 (1-440) x AAD25463 (1-1901)

QY 1 MetLysAlaIleLeuLeuProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
Db 188 ATGAAAGCGATCTTAATCCATTTTATCTCTCTGATTCCGTAAACCCCGCAATCTGCA 247
QY 21 PheAlaGlnSerGluProGlnLeuLeuLeuGluSerValIleValSerArgHisGly 40
Db 248 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGGTGAATTTGTCAGTCTCATGCT 307
QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 308 GTGCGTGTCTCAACACGACGACGCACTGATGAGGATGTCAACCCAGACGCAATGGCCA 367
QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuLeuIleAlaTrpLeu 80
Db 368 ACCTCGCGGTAAACACTGGTGTGGCTGACACCGCGNGGTGGTGAGCTAATCGCCTATCTC 427
QY 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 428 GGACATTACCAACGCGCGCTGTGTAGCCGACGGAATGCTGGCGAAAGGGCTGCGCG 487
QY 101 GlnSerGlyGlnValAlaIleAlaIleAlaAspValAspGluArgLysThrGlyGlu 120
Db 488 CAGTCTGTGTCAGTCCGATTAATCTGATGTGACGAGCGTACCCGTAACACGCGCA 547
QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCCTTCGCGCGCGGCTGGCAGCTGACTGTGCAATAACCGTACATACCCAGCAGATACG 607
QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 608 TCCAGTCCGATCCGCTTAATCTTAATCCTCTAAACACTGCGCTTGCCTCACTGGATAACGCG 667
QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db 668 AACGTGACTGACCGATCTCTAGCAGGCGAGGAGGTCAATGTGACTTACCGGGAT 727
QY 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 728 CGCAACCGCGCTTCGCGAATCGAACCGGTGCTTAATTTTCGCAATCAAACTTGTGC 787
QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 788 CTTAAACGTGAGAAACAGGCGAAAGCTGTTCATTAACGCGACGATTAACCATCGGAATC 847
QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 848 AAGGTGACGCGACATGTCTATTAACCGGTGCGTAACCTCGCATCAATGCTGACG 907
QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Db 908 GAGATATTTCTCTGCAACAGCACAGGGAATCGCGAGCGCGGTGGGGAAGGATCAC 967
QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
Db 968 GATTACACACCATGAGTGGACACCTTGTCTAAGTTGTCATACGCGCAATTTTATTGCTACA 1027
QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 1028 CGCAGCCAGAGTTGCCCGAGCGCGCCACCCGCTTATGATTTGATCAAGACAGCG 1087
QY 301 LeuThrProHisProProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
Db 1088 TTGACGCCCCATCCACCGCAAAACAGCGGTATGTTGATGATTAACCCACTTCAGTGCTG 1147
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 1148 TTTATCGCGGACACGATACTAATCTGGCAATCTCGCGCGCGCACTGGAGCTCAACTGG 1207
QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 1208 ACGTTCCTCGGTACGCGGATTAACACGCGCGCGAGTGTGAGTGTGTTTGAACGCTGG 1267
QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380

Db 1268 CGTCGGCTAACCGATACAGCAGTCAGTTCAGGTTTCGCTGCTTCAGACTTTACAG 1327
QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 1328 CAGATGCGTGATAAACCGCGCTGTCTATTAAATACGCCCGCGGAGAGGTGAACCTGACC 1387
QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1388 CTGCGAGGATGTGAGAGCGGAATGCGCAGGGCATGTGTTGTTGGCAGGTTTACGCA 1447
QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1448 ATCGTGAATGAAGCACGCAATACCGCGCTGCAGTTTG 1483
RESULT 13
ADA19449
ID ADA19449 standard; DNA; 1901 BP.
XX
AC ADA19449;
XX
DT 20-NOV-2003 (first entry)
XX
DE E. coli K12 appA gene encoding phytase.
XX
DS Phytase; ds; gene; phytate; appA gene; animal feed; inorganic phosphate;
KW digestion enhancement; transgenic; thermal tolerance; protease stability.
XX
OS Escherichia coli; strain K12.
XX
FH Key Location/Qualifiers
FT 188..1486
FT /*tag= a
FT /product= "Phytase"
XX
PN US2002136754-A1.
XX
PD 26-SEP-2002.
XX
PF 24-MAY-2001; 2001US-00866379.
XX
PR 13-AUG-1997; 97US-00910798.
PR 01-MAR-1999; 99US-00259214.
PR 13-APR-1999; 99US-00291931.
PR 25-MAY-1999; 99US-00318528.
PR 25-MAY-2000; 2000US-00580515.
XX
PA (SHOR/) SHORT J M.
PA (KRET/) KRETZ K A.
PA (GRAY/) GRAY K A.
PA (BART/) BARTON N R.
PA (GARR/) GARRETT J B.
PA (DONO/) O' DONOGHUE E.
PA (MATH/) MATHUR E J.
XX
PI Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB;
PI O' Donoghue E, Mathur EJ;
XX
DR WPI; 2003-040002/03.
DR P-PSDB; ADA19450.
XX
XX Isolated Escherichia coli polynucleotide encoding a modified phytase
PT enzyme, useful in the production of animal feed, for improving the
PT nutritional value of phytate-containing feedstuff and for enhancing
PT digestion in humans and animals.
XX
PS Claim 42; Fig 7; 62pp; English.
XX
CC The invention relates to an isolated Escherichia coli polynucleotide
CC encoding a phytase enzyme appearing as ADA19450 and having amino acids
CC modification from W68E, Q84W, A95P, K97C, S168E, R180Y, N262C or Y277D.
CC Also included the E. coli appA gene ADA19449 (or an oligonucleotide
CC derived from it) or its mutant sequence ADA19452, expression vectors,
CC

CC host cells, a method of improving nutritional value of a phytate-
CC containing foodstuff by contacting the phytate-containing foodstuff with
CC a pure phytase enzyme mutated phytase (where the phytase enzyme catalyzes
CC the liberation of inorganic phosphate from the phytate in the phytate-
CC containing foodstuff), a method to produce an animal feed containing a
CC microbial phytase (comprising culturing the plant cell, plant part or
CC plant under conditions where the nucleotide sequence is expressed and
CC converting the plant cells, plants or plant into a composition for animal
CC feed), a feed composition for animals (comprising the plant seeds, plant
CC cells, plant parts or plants in admixture with a phytase-containing
CC foodstuff), a method to treat a human or an animal able to benefit from
CC digestive enhancement by the activity of an exogenous phytase enzymes
CC comprising administering to the human or animal the plant seed, plant
CC cells, plant parts or plants of a transgenic plant which is modified to
CC contain an expression system which expresses a nucleotide sequence
CC encoding a phytase enzyme, a transgenic non-human organism whose genome
CC comprising a heterologous nucleic acid sequence encoding a polypeptide
CC having phytase activity. The phytase enzyme is useful for improving the
CC nutritional value of phytate-containing foodstuff, in the production of
CC animal feed and for enhancing digestion in humans and animals. The
CC invented method improves thermal tolerance and protease stability. It
CC also improves the feeding value of phytate rich ingredients. The present
CC sequence represents the E. coli K12 appA gene encoding wild-type phytase.
XX
SQ Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 7,83e-196 Length: 1901
Score: 2235.00 Matches: 430
Percent Similarity: 99.5% Conservative: 0
Best Local Similarity: 99.5% Mismatches: 2
Query Match: 97.1% Indels: 0
DB: 8 Gaps: 0

US-10-601-319-2 (1-440) x ADA19449 (1-1901)

Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 188 ATGAAAGCGATCTTAATCCCAATTTTATCTTCTGATTCGTTAACCCCGCAATCTGCA 247
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db 248 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGCGAAAGTGTGGTATGTGTCAGTCTCATGCT 307
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
Db 308 GTGCGTCTCCAAACCAAGGCCACGCAACTGATGCGAGATGTCAACCCAGACGATGCGCCA 367
Qy 61 ThrTrpProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db 368 ACCTGGCCGGTAAACCTGGGTGGCTGACACCGCGNGGTGGTGAGCTAATCGCTATCTC 427
Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 428 GGACATTACCAAGCCAGCGTCTGTAGCCGACGAGATTGTGGCGAAAGAGGCGTGGCCG 487
Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 488 CAGTCTCGTCAGGTCGCGATATTATGCTGATGTGCACGAGCGTACCCGTAATAACAGCGGAA 547
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCCTTCGCCCGCGGCTGGCACCTGATGTGCAATAACCGTACATACCCAGGCGAGATACG 607
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 608 TCCAGTCCCGATCGTTATTAATCTCTCTAAACCTGGCGTGTGGCAACTGGATAACGCG 667
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db 668 AACGTGACTACGCGATCCCTCAGACGGCGAGGGGTCAATGCTGATTTTACCGGCAT 727
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200

Db 728 CGGCAAAACGGCGTTTCGCGACTGGAAACGGGTGCTTAATTTCCGCAATCAACTTGTGC 787
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 788 CTTAAACGTCGAGAAACAGGACGAAAGCTGTTCAATTAACGAGGATATCCATCGGAATC 847
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 848 AAGGTGAGCGCGACAAATGTCATTAAACGGGTGCGTAAGCCTCGCATCAATGCTGACG 907
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
Db 908 GAGATATTCTCTCGCAACAGCACAGGAAATGCCGAGCGCGGGTGGGAGGATCACC 967
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 968 GATTACACAGCAGTGAACACCTTGCTAGTTTGCATAACGCGCAATTTTATTGCTACAA 1027
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 1028 CGCACGCCAGAGGTGCGCGACGCGCCACCCCGTTATTAGATTGATCAAGACAGCG 1087
Qy 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 1088 TTGACGCGCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTCTG 1147
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 1148 TTTATCGCGGACACGATACTAATCTGGCAAAATCTCGCGCGGCGCACTGGAGGCTCAACTG 1207
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 1208 AGCTTTCGCGGTGAGCGCGATACCGCGCGAGGTGGTGAACCTGTTGAACGCTGG 1267
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1268 CGTCGCGTAAAGCATTAACAGCCAGTGGATTTCAGTTTCGCTGCTTCCAGACTTTACAG 1327
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 1328 CAGATCGGTGATAAAACGCGCTGTCTATTAAATACGCGCGCGAGAGGTGAAACTGACC 1387
Qy 401 LeuAlaGlyCysGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1388 CTGGCAGGATGGAAGAGCGAAATGCGAGGCGATGTGTCTGTGCGAGGTTTACGCAA 1447
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1448 ATCGTGAATGAGCAGCATACCGCGCTGCGATTG 1483
RESULT 14
ADC87744
ID ADC87744 standard; DNA; 1901 BP.
XX
AC ADC87744;
XX
DT 01-JAN-2004 (first entry)
XX
DE DNA encoding Escherichia coli appA phytase.
XX
KW Phytase: food supplement; enzyme delivery matrix; soybean meal;
KW thermotolerance; thermostability; kernel; phytate;
KW myo-inositol-hexaphosphate; inositol; inorganic phosphate;
KW thermotolerant; feed value; digestion; gene; ds; appA.
XX
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT CDS 188..1486
FT /*tag= a
FT /product= "AppA phytase"
FT /transl_except= (pos:401..403,aa:Arg)

|||||
1388 CTGCAGATGTGAAGAGCGAAATGCCAGGCGATGTTCTGTGGCAGGTTTACGCAA 1447
Qy 421 ILeValAsnGluAlaArgIleProAlaCysSerLeu 432
|||||
Db 1448 ATCGTGAATGAAGCAGCAGCATACCGCGCTGCAGTTTG 1483
RESULT 15
ADO50297
ID ADO50297 standard; DNA; 1901 BP.
XX AC ADO50297;
XX 29-JUL-2004 (first entry)
XX DE Escherichia coli K12 appA phytase DNA.
XX KW AppA phytase; bacteria; thermal tolerance; protease stability; foodstuff;
XX KW phytate; animal feed; fish feed; dough; baking; gene; ds.
XX OS Escherichia coli; K12.
XX FH Key Location/Qualifiers
XX CDS 188..1486
FT /*tag= a
FT /product= "AppA phytase"
FT 401..403
FT unsure /*tag= b
FT /*note= "Encodes Arg"
XX US2004091968-A1.
XX 13-MAY-2004.
XX 20-JUN-2003; 2003US-00601319.
XX 13-AUG-1997; 97US-00910798.
XX 01-MAR-1999; 99US-00259214.
XX 13-APR-1999; 99US-00291931.
XX 25-MAY-1999; 99US-00318528.
XX 25-MAY-2000; 2000US-00580515.
XX 24-MAY-2001; 2001US-00866379.
XX (SHOR/) SHORT J M.
XX (KRET/) KRETZ K.
XX (GRAY/) GRAY K A.
XX (BART/) BARTON N R.
XX (GARR/) GARRETT J B.
XX (ODON/) O'DONOGHUE E.
XX (MATH/) MATHER E J.
XX Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
XX Mather EJ;
XX WPI: 2004-374952/35.
XX P-PSDB; ADO50298.
XX Producing phytase, involves providing nucleic acid derived from bacteria
XX encoding polypeptide having phytase activity, and expressing nucleic acid
XX in yeast.
XX Disclosure; SEQ ID NO 7; 74pp; English.
XX The invention relates to a method for producing phytase that involves
XX providing a nucleic acid encoding phytase derived from a bacteria, and
XX expressing the nucleic acid in a yeast under conditions that allow
XX expression of the enzyme in the yeast. The invention also relates to
XX modified phytase enzyme which has improved thermal tolerance and protease
XX stability at low pH. The phytase enzyme can be used in foodstuffs to
XX improve the feeding value of phytate rich ingredients, and in diet of
XX numerous animals including mammals, fowls and fishes, commercially
XX significant mammals such as pigs, goats, laboratory rodents, commercially
XX significant avian species such as chicken, ducks, doves, parrot, etc.,

CC commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC etc., in dough making and baking, in dietary aids for animals. The method
CC provides easy manufacture of the active ingredient loaded biocompatible
CC composition, higher yields and loading efficiency. The phytase
CC incorporated in to the dietary aid is safe for animals. The present
CC sequence is Escherichia coli K12 appA phytase DNA.
XX
SQ Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 0 U; 1 Other;
Alignment Scores:
Pred. No.: 7.83e-196 Length: 1901
Score: 2235.00 Matches: 430
Percent Similarity: 99.5% Conservatives: 0
Best Local Similarity: 99.5% Mismatches: 2
Query Match: 97.1% Indels: 0
DB: 12 Gaps: 0
US-10-601-319-2 (1-440) x ADO50297 (1-1901)
Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 188 ATGAAAGCGATCTTAATCCCATTTTATCTCTTCTGATTCCGTTAACCCCGCAATCGCA 247
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db 248 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAGTGTGGTATTGTCTGCTGTCTATGGT 307
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 308 GTGCGTGTCTCAACCAAGCCAGCCAGCAACTGATGTCAGAGATGTCAACCCAGACGCATGGCCA 367
Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
Db 368 ACCTGGCCGGTAAACTGGGTTGGCTGACACCCCGNGGTGGTGAGCTAATTCGCTATCTC 427
Qy 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 428 GGACATTTCACACGCCAGCGGTCTGGTAGCCGCGATTGCTGGCGAAAGAGGCTGCCCG 487
Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 488 CAGTCTGGTCAGGTGCGGATTATTGCTGATGTGTCAGAGCGGTACCCGTAACAGCGCGAA 547
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCCTTCGCCCGCGGCTGGCACCTGACTGTGCAATAAACCGTATACCCAGGCGAGATACG 607
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 608 TCCAGTCCCGATCCGTTATTATTCCTCTAANAACCTGGGTTGGCAACTGGATACGCG 667
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db 668 AACGTGACTGACGCGATCTCTCAGCAGGCGAGGAGGTCAATTGCTGACTTTTACCGGCGAT 727
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 728 CGCAACACCGCGTTTCGCAACTGGAAACCGGTGCTTAATTTTCCGCAATCAAACTTGTGC 787
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 788 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCTAATACGACGAGCATATACCATCGGAATC 847
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 848 AAGGTGAGCGCGCACAAATGCTCATTAACCGGTGGGTAAAGCTCGCATCAATGCTGACG 907
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Db 908 GAGATATTCTCTCTGCAACAGCACAGGGAATCCCGAGCCGGGTGGGGAAGATACCC 967
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280

Db 968 GATTCACACAGTGGAAACACCTTGCTAAGTTTGCATAAACGCGCAATTTTATTGCTACAA 1027
Qy ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleuMetAlaAla 300
Db |||||
1028 CGCAGCCAGAGGTGGCCGCGCAGCCGCCGCCCGTATTAGATTGATCAAGACAGCG 1087
Qy LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db |||||
1088 TTGACGCCCATCCACCGCAAAACAGGCGTATGTTGTGACATTACCCACTTCAGTGTCTG 1147
Qy PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db |||||
1148 TTTATCGCGGACACGATACATACTGGCAATCTCGCGCGGCACTGGAGCTCAACTGG 1207
Qy ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db |||||
1208 ACGCTTCGCGTCCAGCCGCGATAACACGCGCGCAGGTGTGTGAAGTGTGTTGACGCTGG 1267
Qy ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrIleuGln 380
Db |||||
1268 CGTCGGCTAAGCGATAACAGCCAGTGGATTTCAGGTTTCGCTGGTCTCCAGACTTTACAG 1327
Qy GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db |||||
1328 CAGATGCGTGATAAAACCGCGTGTCAATTAATACGCGCGCGGAGAGGTGAAACTGACC 1387
Qy LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db |||||
1388 CTGGCAGGATGTGAAGACGGAATGCGCAGGCGCATGTGTTGTTGGCAGGTTTACGCCAA 1447
Qy IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db |||||
1448 ATCGTGAATGAAGCAGCAGCATACCGGCGTGCAGTTTG 1483

Search completed: June 14, 2006, 10:57:59
Job time : 591.615 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 14, 2006, 10:58:44 ; Search time 5255.78 Seconds
(without alignments)
8030.264 Million cell updates/sec

Title: US-10-601-319-2

Perfect score: 2302
Sequence: 1 MMAILPFLSLILPTQSA.....IVNEARIPACSLRSHHHHH 440

Scoring table:

BLOsum62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abes/ABSSWEB spool/US10601319/runat 12062006 180050 27480/app query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs06h
-USER=US10601319 @CCN 1.1 3991 @runat 12062006 180050 27480 -NCPU=6 -ICPU=3
-NO.MMAP -NRG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb env.*
2: gb pat.*
3: gb ph.*
4: gb pl.*
5: gb pr.*
6: gb ro.*
7: gb sta.*
8: gb sy.*
9: gb un.*
10: gb vi.*
11: gb ov.*
12: gb htg.*
13: gb in.*
14: gb om.*
15: gb pa.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2302	100.0	1323	2	ARI08133 Sequence
2	2302	100.0	1323	2	ARI27818 Sequence
3	2302	100.0	1323	2	ARI30956 Sequence

4	2302	100.0	1323	2	AR492966	AR492966 Sequence
5	2302	100.0	1323	2	AR636185	AR636185 Sequence
6	2302	100.0	1323	2	AX052713	AX052713 Sequence
7	2302	100.0	1323	2	AX338966	AX338966 Sequence
8	2302	100.0	1323	2	AX356566	AX356566 Sequence
9	2302	100.0	1323	2	AX356572	AX356572 Sequence
10	2302	100.0	1323	2	AX356572	AX356572 Sequence
11	2302	100.0	1323	2	AX356572	AX356572 Sequence
12	2302	100.0	1323	2	AX356572	AX356572 Sequence
13	2302	100.0	1323	2	AX356572	AX356572 Sequence
14	2302	100.0	1323	2	AX356572	AX356572 Sequence
15	2302	100.0	1323	2	AX356572	AX356572 Sequence
16	2302	100.0	1323	2	AX356572	AX356572 Sequence
17	2302	100.0	1323	2	AX356572	AX356572 Sequence
18	2302	100.0	1323	2	AX356572	AX356572 Sequence
19	2302	100.0	1323	2	AX356572	AX356572 Sequence
20	2302	100.0	1323	2	AX356572	AX356572 Sequence
21	2302	100.0	1323	2	AX356572	AX356572 Sequence
22	2302	100.0	1323	2	AX356572	AX356572 Sequence
23	2302	100.0	1323	2	AX356572	AX356572 Sequence
24	2302	100.0	1323	2	AX356572	AX356572 Sequence
25	2302	100.0	1323	2	AX356572	AX356572 Sequence
26	2302	100.0	1323	2	AX356572	AX356572 Sequence
27	2302	100.0	1323	2	AX356572	AX356572 Sequence
28	2302	100.0	1323	2	AX356572	AX356572 Sequence
29	2302	100.0	1323	2	AX356572	AX356572 Sequence
30	2302	100.0	1323	2	AX356572	AX356572 Sequence
31	2302	100.0	1323	2	AX356572	AX356572 Sequence
32	2302	100.0	1323	2	AX356572	AX356572 Sequence
33	2302	100.0	1323	2	AX356572	AX356572 Sequence
34	2302	100.0	1323	2	AX356572	AX356572 Sequence
35	2302	100.0	1323	2	AX356572	AX356572 Sequence
36	2302	100.0	1323	2	AX356572	AX356572 Sequence
37	2302	100.0	1323	2	AX356572	AX356572 Sequence
38	2302	100.0	1323	2	AX356572	AX356572 Sequence
39	2302	100.0	1323	2	AX356572	AX356572 Sequence
40	2302	100.0	1323	2	AX356572	AX356572 Sequence
41	2302	100.0	1323	2	AX356572	AX356572 Sequence
42	2302	100.0	1323	2	AX356572	AX356572 Sequence
43	2302	100.0	1323	2	AX356572	AX356572 Sequence
44	2302	100.0	1323	2	AX356572	AX356572 Sequence
45	2302	100.0	1323	2	AX356572	AX356572 Sequence

ALIGNMENTS

RESULT 1
ARI08133
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
ORIGIN

Sequence 1 from patent US 6110719.
ARI08133
ARI08133.1 GI:12823620
Unknown.
Unclassified.
1 (bases 1 to 1323)
Kretz, K.
Phytase
Patent: US 6110719-A 1 29-AUG-2000;
Location/Qualifiers
1..1323
/organism="unknown"
/mol_type="unassigned DNA"

Alignment Scores:
Pred. No.: 8.92e-205 Length: 1323
Score: 2302.00 Matches: 440
Percent Similarity: 100.0% Conservativeness: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

```
US-10-601-319-2 (1-440) x AR108133 (1-1323)
QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAAGCGATCTTAATCCCATTTTATCTCTCTGATTCCGTTAAACCCCGCAATCTGCA 60
QY 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db 61 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGAAGTGTGGTGAATTTGTCAGTCTGTCATGGT 120
QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTTPPro 60
Db 121 GTGGCTCTCAACCAAGGCCACCACTGATGAGGATGTACCCCGAGAGCGATGGCCA 180
QY 61 ThrTrpProValLysLeuGlyTTPLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db 181 ACCTGGCCGGTAAAACTGGGTGGCTGACACCGCGNGGTGGTGAAGCTAATCGCCTATCTC 240
QY 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 241 GGACATTACCAACCGCAGCGCTGTGTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG 300
QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 301 CAGTCTGGTCAGGTCCGATTAATTCCTTAATAACTGGCGTGTGACAGCGTACCCGTAAACAGGCGAA 360
QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCCTTCGCGCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGAGATACG 420
QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCCAGTCCCGATCCGTTATTTAATTCCTTAATAACTGGCGTGTGACAGCGTACCCGTAAACAGGCG 480
QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGACTGACGCGATCTCTGACGAGGCGAGGAGGTCAATTCGTGACTTTACCGGCAAT 540
QY 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 541 CGGCAAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCGCGCAATCAAACTTGTC 600
QY 201 LeuLysArgGlnLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTCTAATACGAGGCGATTAACCATCGGAATC 660
QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGAGCGCGCAATGCTCAATTAACCGGTGCGGTAAAGCTCGCATCAATGCTGACG 720
QY 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTTPGlyArgIleThr 260
Db 721 GAGATAATTTCTCTGCAACAGCACAGCGAATGCGGAGCGGGGTGGGGAAGGATCAC 780
QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 781 GATTACACACGAGTGGAAACACTTCTTAAGTTTGCATTAACCGCGCAATTTTATTTGCTACAA 840
QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 841 CGACGCGCAGAGGTTCGCGCAGCGCGGCCACCCCGTTATGATTTGATCATGGCAGCG 900
QY 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 901 TTGACGCGCCCATCCACGCAAAACAGCGCTATGCTGACATTAACCACTTCAGTACTG 960
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 961 TTTATTCGCGGACACGATACTAATCTGGGAAATCTCGCGCGGCGACCTGGAGCTCAACTGG 1020
QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTTP 360
Db 1021 ACCTTCCCGGTGAGCGGATAACACCGCCAGGTGGTGAACCTGGTGTGAAACGCTGG 1080
QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1081 CGTCGGCTAAGCGATAACAGCCAGTGGATTTCAGGTTTCGCTGCTTCCAGAGCTTTACAG 1140
QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 1141 CAGATCGTGTATAAACCGCGCTGTCAATTAATACCGCCCGGAGAGGTGAAACTGACC 1200
QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1201 CTGGCAGGATGTGAAGCGCAATATGCGAGGCGATGTTCTGTTGGCAGAGTTTACGCAA 1260
QY 421 IleValAsnGluAlaAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
Db 1261 ATCGTGAATGAAGCAGCATACCGCGGTGCAGTTTGAGATCTCATACCATCATCATCAC 1320
RESULT 2
AR127818 LOCUS 1323 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6183740.
ACCESSION AR127818
VERSION AR127818.1 GI:14115480
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1323)
AUTHORS Short, J.M. and Kretz, K.A.
TITLE Recombinant bacterial phytases and uses thereof
JOURNAL Patent: US 6183740-A 1 06-FEB-2001;
FEATURES Location/Qualifiers
source 1..1323
/mol_type="unassigned DNA"
ORIGIN
Alignment Scores:
Pred. No.: 8,92e-205 Length: 1323
Score: 2302.00 Matches: 440
Percent Similarity: 100.0% Conservat: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0
US-10-601-319-2 (1-440) x AR127818 (1-1323)
QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAAGCGATCTTAATCCCATTTTATCTCTCTGATTCCGTTAAACCCCGCAATCTGCA 60
QY 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db 61 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGAAGTGTGGTGAATTTGTCAGTCTGTCATGGT 120
QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTTPPro 60
Db 121 GTGGCTCTCAACCAAGGCCACCACTGATGAGGATGTACCCCGAGAGCGATGGCCA 180
QY 61 ThrTrpProValLysLeuGlyTTPLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db 181 ACCTGGCCGGTAAAACTGGGTGGCTGACACCGCGNGGTGGTGAAGCTAATCGCCTATCTC 240
QY 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 241 GGACATTACCAACCGCAGCGCTGTGTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCCG 300
QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 301 CAGTCTGGTCAGGTCCGATTAATTCCTTAATAACTGGCGTGTGACAGCGTACCCGTAAACAGGCGAA 360
QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
```

```
Db 361 GCCTTCGCGCGGCTGGCACTGCTGCAATAACCGTACATACCAGCGCATACG 420
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCCAGTCCGATCCGTTATTAATCTCTAANAACCTGGCGTTGGCACTGGATACCGG 480
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGACTGACGCGATCCTCAGCAGGCGCAGAGGCTCAATTCGTGACTTTACCGGCAT 540
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 541 CGCAAAACGGCGTTTCGGGAACCTGAACCGGTGCTTAATTTTCGCCAATCAAACTTTGTC 600
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAAACGTGAGAAACAGACGAAAGCTGTTCAATTAACGCGGCAATTAACCATCGGAATC 660
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGAGCGCGCAATGCTCAATTAACCGGTGCGGTAAAGCCTCGCATCAATGCTGACG 720
Qy 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
Db 721 GAGATATTTCTCTGCAACAAGCACAGGAAATGCCGAGCGCGGGTGGGGAAGGATCACC 780
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 781 GATTACACACAGTGGAAACCTTGTCTAAGTTTGGATTAACCGGCAATTTATTTGCTACAA 840
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 841 CGCACGCGCAGAGTTGCCCGCAGCGCCACCCCGTTATTTGGATTTGATCATGGCAGCG 900
Qy 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 901 TTGACGCGCCCATCCACCGCAAAAACAGCGGTATGCTGTGACATTAACCCACATTCAGTACTG 960
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeuGluLeuAsnTrp 340
Db 961 TTTATTCGCGGACACGATACTAATCTGGCAATCTCGCGCGGCACTGGAGCTCAACTGG 1020
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 1021 ACGCTTCCCGTCAAGCGGATAACAGCGCGCCAGTGGTGAACCTGGTTTGAACGCTGG 1080
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1081 CGTCGGCTAAGCGATAACAGCGCAGTGGATTTCAGGTTTCAGGTTTCAGACTTTACAG 1140
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 1141 CAGATGGGTGATAAACCCTGCTCATTAATACGCCCGCGGAGAGGTGAACACTGACC 1200
Qy 401 LeuAlaGlyCysGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1201 CTGCGAGGATGTGAAGACGGAATCGCGAGGCAATGTTGCTGTTGGCAGGTTTACGCAA 1260
Qy 421 IleValLeuGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
Db 1261 ATCGTGAATGAACGACGATACCGCGCGTGGATTGTAGATCTCATCACCATCACCATCAC 1320
```

RESULT 3

LOCUS AR130956 1323 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6190897.

ACCESSION AR130956

VERSION AR130956.1 GI:14119281

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 1323)

AUTHORS Kretz, K.
TITLE Phytase
JOURNAL Patent: US 6190897-A 1 20-FEB-2001;
FEATURES Location/Qualifiers
source 1..1323
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:

Pred. No.: 8.92e-205 Length: 1323
Score: 2302.00 Matches: 440
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
Dbs: 2 Gaps: 0

US-10-601-319-2 (1-440) x AR130956 (1-1323)

```
Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAGCGATCTTAATCCCATTTTATCTTCTGATTCGGTTAACCCCGCAATCTGCA 60
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db 61 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAAGTGTGGTATTGTCTCAGTCGTCACTGT 120
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 121 GTGCGTCTCTCAACCAAGGCCACGCACTGATGATGAGGATGTCAACCCAGCAGCATGGCCA 180
Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db 181 ACCTGGCGGTAAACATGGGTGGCTGACACCGCGGGTGGTGGAGCTAATCGCTATCTC 240
Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 241 GGACATTTACCAACGCCAGCGCTCTGGTAGCGGACGGATTGCTGGCGMAAAGGGCTGCCCG 300
Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 301 CAGTCTGGTCAAGTTCGCGGATTTATTTGCTGATGTCCACGAGCGTACCCGTTAAACAGCGCAA 360
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCCTTCGCGCGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGCGCAGATACG 420
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCCAGTCCCGATCCGTTATTTAATCCTCTAANAACCTGGCGTTTGCCAACTCGAATAACGCG 480
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGACTGACGCGATCCTCAGCAGGCGCAGAGGCTCAATTTGCTGACTTTACCGGCAT 540
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 541 CGGCAAAACGGCGTTTCGGGAACCTGAACCGGTGCTTAATTTTCGCCAATCAAACTTTGTC 600
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAAACGTGAGAAACAGACGAAAGCTGTTCAATTAACGCGGCAATTAACCATCGGAATC 660
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGAGCGCGCAATGCTCAATTAACCGGTGCGGTAAAGCCTCGCATCAATGCTGACG 720
Qy 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
Db 721 GAGATATTTCTCTGCAACAAGCACAGGAAATGCCGAGCGCGGGTGGGGAAGGATCACC 780
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
```

```
Db      781  |||||GATTACACAGTGGAAACACCTTGCTAAGTTTGCTAATACGCGCAATTTATTTGCTACAA      840
Qy      281  ArgThrProGluValAlaAArgSerArgAlaThrProLeuLeuLeuLeuLeuMetAlaAla      300
Db      841  CGCAGCGCAGAGGTGCGCGCAGCGCGCCACCCCGTTATTGGATTGATCATGGCAGCG      900
Qy      301  LeuThrProHisProProGlnLysGlnAlaTyGlyValThrLeuProThrSerValLeu      320
Db      901  TTGACGCCCATCCACCGCAAAACAGCGGTATGGTGACATTACCACCTTCAGTACTG      960
Qy      321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp      340
Db      961  TTTATTGCCGACAGGATACTAATCTGGCAAACTCGCGCGCGCACTGGAGCTCAACTGG      1020
Qy      341  ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp      360
Db      1021  ACGCTTCCCGGTACGCGGATAACACGCGCCAGGTGGTGAACCTGGTTTGAACGTGG      1080
Qy      361  ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln      380
Db      1081  CGTCGGCTAAGCGGATAACAGCGAGTGAATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG      1140
Qy      381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr      400
Db      1141  CAGATGGCTGATAAACCGCGCTGTCTAATTAATACGCGCGCGGAGAGGTGAACCTGACC      1200
Qy      401  LeuAlaGlyCysGluGluAArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln      420
Db      1201  CTGGCAGGATGTGAAGAGCGAAATCGCAGGGCATGTGTTGTTGGCAGGTTTACGCAA      1260
Qy      421  IleValAsnGluAlaAArgIleProAlaCysSerLeuArgSerHisHisHisHisHis      440
Db      1261  ATCGTGAATGAAGCACGATACCGCGCTGCAGTTTGAGATCTCATCACCATTCACTAC      1320

RESULT 4
AR492966 LOCUS AR492966 1323 bp DNA linear PAT 15-MAY-2004
DEFINITION Sequence 1 from patent US 6720014.
ACCESSION AR492966
VERSION AR492966.1 GI:47264255
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1323)
AUTHORS Short,J.M. and Kretz,K.A.
TITLE Phytase-containing foodstuffs and methods of making and using them
JOURNAL Patent: US 6720014-A 1 13-APR-2004;
Diversa Corporation; San Diego, CA
FEATURES
source
location/Qualifiers
1..1323
/mol_type="genomic DNA"
/organism="unknown"

ORIGIN

Alignment Scores:
Pred. No.: 8.92e-205 Length: 1323
Score: 2302.00 Matches: 440
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-601-319-2 (1-440) x AR492966 (1-1323)
Qy      1  MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla      20
Db      1  ATGAAACGATCTTAATCCCAATTTTATCTCTGATTCGTTAACCCCGCGCAATCTGCA      60
Qy      21  PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly      40
Db      61  TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGMAAGTGTGGTATTGTCAGTCGTCATGGT      120
```

```
Qy      41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro      60
Db      121  GTCGGTGCTCAACCAAGGCCACGCAACTGATCAGAGATGTACCCCGACGCGATGGCCA      180
Qy      61  ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyLeu      80
Db      181  ACCTGGCCGGTAAAACTGGGTTGGCTGACACCGCGNGGTGGTGAGCTAATCGCTATCTC      240
Qy      81  GlyHisTyGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro      100
Db      241  GGACATTATACCAAGCCAGCGCTCTGGTAGCCGCGATTGCTGGCGAAAAAGGGCTGCCG      300
Qy      101  GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu      120
Db      301  CAGTCTGGTCAAGTCGCGATTATTGTGTGATGTGACGAGCGTACCCGTAACAAAGGCGAA      360
Qy      121  AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr      140
Db      361  GCCTTCGCCCGCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGCGCAGATACG      420
Qy      141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla      160
Db      421  TCCAGTCCCGATCCGTTATTAAATCCTCTAAAAAAGTGGCGTTTGGCAACTGGATAACGCG      480
Qy      161  AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis      180
Db      481  AACGTGACTGACCGGATCCTCAGCAGGCGAGGAGGTCAATTGTCTGACTTTTACCGGGCAT      540
Qy      181  ArgGlnThrAlaPheArgGluLeuArgValLeuAsnPheProGlnSerAsnLeuCys      200
Db      541  CGGCAAAACGGGTTTCGCGNACTGGAAACGGGTCTTAATTTTCGCAATCAAACTTGTGC      600
Qy      201  LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu      220
Db      601  CTTAAACGTCGAGAAACAGGACGAAAGCTGTTCTATTAAACGCGAGCATTTACCATCGAACTC      660
Qy      221  LysValSerAlaAspAsnValSerLeuThrGlyValValSerLeuAlaSerMetLeuThr      240
Db      661  AAGTGAGCGCCGACAAATGTCTCATTAACCGGTGGGTAAAGCTTCGATCAATGCTGACG      720
Qy      241  GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr      260
Db      721  GAGATATTCTCTGCAACACAGCACAGGGAATGTCGAGGCGCGGGTGGGGAAGATCACC      780
Qy      261  AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyLeuLeuGln      280
Db      781  GATTCACACAGTGGAAACACTTGTCTAAGTTTGCATAACGCGCAATTTTATTATTGCTACAA      840
Qy      281  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla      300
Db      841  CGCAGCCGAGAGGTTCGCCGACGCGCGCCACCCCGTTATTGGATTGATCATGGCAGCG      900
Qy      301  LeuThrProHisProProGlnLysGlnAlaTyGlyValThrLeuProThrSerValLeu      320
Db      901  TTGACGCCCATCCACCGCAAAACAGCGGTATGGTGACATTACCACCTTCAGTACTG      960
Qy      321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuLeuAsnTrp      340
Db      961  TTTATTGCCGACAGGATACTAATCTGGCAAACTCGCGCGCGCACTGGAGCTCAACTGG      1020
Qy      341  ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp      360
Db      1021  ACGCTTCCCGGTACGCGGATAACACGCGCGCGAGGTGGTGAACCTGGTTTGAACCGCTGG      1080
Qy      361  ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln      380
Db      1081  CGTCGGCTAAGCGGATAACAGCGAGTGAATTCAGGTTTCGCTGGTCTTCCAGACTTTACAG      1140
Qy      381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr      400
Db      1141  CAGATCGTGATAAACCGCGCTGTCTAATTAATACGCGCGCGGAGAGGTGAACCTGACC      1200
```

QY 401 LeuAlaGlyCysGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 Db 1201 CTGGCAGGATGTGAAGAGCGAAATCGCAGGCGATGTGTCTGTCGCGAGGTTTACGCNA 1260

QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
 Db 1261 ATCGTGATGATGACGACGATACCGCGGTGCGATTTGAGATCTCATCACCATCACCATCAC 1320

RESULT 5
 AR636185
 LOCUS AR636185 1323 bp DNA linear PAT 20-APR-2005
 DEFINITION Sequence 1 from patent US 6855365.
 ACCESSION AR636185
 VERSION AR636185.1 GI:62768868
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 1323)
 AUTHORS Short, J.M., Kretz, K.A., Gray, K.A., Barton, N.R., Garrett, J.B.,
 O'Donoghue, E. and Mathur, E.J.
 TITLE Recombinant bacterial phytases and uses thereof
 JOURNAL Patent: US 6855365-A 1 15-FEB-2005;
 Diversa Corporation; San Diego, CA
 FEATURES
 Location/Qualifiers
 source 1..1323
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN
 Alignment Scores:
 Pred. No.: 8.92e-205 Length: 1323
 Score: 2302.00 Matches: 440
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 2 Gaps: 0

US-10-601-319-2 (1-440) x AR636185 (1-1323)

QY 1 MetLysAlaIleLeuLeuProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
 Db 1 ATGAAGCGATCTTAATCCCAATTTTATCTCTCTGATTCGGTAAACCCCGCAATCTGCA 60

QY 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40
 Db 61 TTCGTCAGATGAGCGGAGCTGAAGCTGGAAGTGTGGTGTGATGTGTCGTCGTATGGT 120

QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 Db 121 GTGGGTGTCTCAACCAAGGCCACGCACTGATGTCAGGATGTCAACCCAGACGCGATGGCA 180

QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaThrLeu 80
 Db 181 ACCGTGGCGGTAAACATGGGTGTGCTGACACCGCGGTGGTGTGATTAATCGCCTATCTC 240

QY 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 Db 241 GGACATTAACCAACGCCAGCGTCTGTGACCGACCGATGCTGGCGAAAGGGGTGCCCG 300

QY 101 GlnSerGlyGlnValAlaIleAlaAspValAspValAspGluArgThrArgLysThrGlyGlu 120
 Db 301 CAGTCTGGTCAGGTCCGCGATTTATGCTGATGTCAGCAGCGGTACCCGTAAACAGGCGAA 360

QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 Db 361 GCCTTCGCGCGCGGTGGACCTGTGATCTGTGCAATAACCGTATACCCAGCGCAGATACG 420

QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 Db 421 TCCAGTCCCGATCGGTATTTAATCTCTAANAACATGCGCGTTGCCAACTGGATTAACGCG 480

QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180

Db 481 AACGTGACTGACGCGATCTCTCAGCAGGCGCAGGCGTCAAATTGCTGACTTTACCGGCAT 540

QY 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
 Db 541 CGGCANAACCGCGTTCGCGAACTGGAACGGGTGCTTAATTTTCGCAATCAAACTTTGTC 600

QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 Db 601 CTAAACGTGAGAACACAGGACGAAAGCTGTTCAATTAACGCGAGCATTACCATCGGAATC 660

QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 Db 661 AAGGTGAGCGCGCAATGTCTCAATTAACCGGTGCGTAAGCTCGCATCAATGCTGACG 720

QY 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
 Db 721 GAGATATTCTCTGCAACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 780

QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyLeuLeuGln 280
 Db 781 GATTACACCACTGGAACACCTTCTAAGTTTGCATTAACGCGCAATTTATTTTGTGTACA 840

QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 Db 841 CGCACGCGCAGAGGTGCGCGCAGCGCCACCCGCTTATTTGGATTTGATCATGCGACGCG 900

QY 301 LeuThrProHisProProGlnLysGlnAlaTyGlyValThrLeuProThrSerValLeu 320
 Db 901 TTGACGCGCCCATCCACCGCAAAACAGCGGTATGCTGTGACATTTACCCACTTCAGTACTG 960

QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 Db 961 TTTATTGCGGACACGATCTAATCTGGCAATCTCGCGCGCGACTGGAGCTCAACTGG 1020

QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 Db 1021 ACCTTCCCGTCAGCGGATTAACACGCGCGGTGCTGAATCTGCTGGTGTGTAACGCTGG 1080

QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 Db 1081 CGTCGGCTAAGCGGATAACACCGCGTGTCAATTAATACGCGCGCGGAGGAGGTTTACAG 1140

QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 Db 1141 CAGATGGTGATAAAACCGCGCTGTCAATTAATACGCGCGCGGAGGAGGTTTACAG 1200

QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 Db 1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGCGATGTGTTGTTGCGCAGGTTTACGCA 1260

QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
 Db 1261 ATCGTGAATGAACACGATACCGCGGTGCGATTTGAGATCTCATCACCATCACCATCAC 1320

RESULT 6
 AX052713
 LOCUS AX052713 1323 bp DNA linear PAT 12-JAN-2001
 DEFINITION Sequence 1 from Patent WO0071728.
 ACCESSION AX052713
 VERSION AX052713.1 GI:12226903
 KEYWORDS
 SOURCE Escherichia coli
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 Enterobacteriaceae; Escherichia.
 REFERENCE 1
 AUTHORS Short, J.M. and Kretz, K.A.
 TITLE Recombinant bacterial phytases and uses thereof
 JOURNAL Patent: WO 0071728-A 1 30-NOV-2000;
 Diversa Corporation (US)
 FEATURES
 Location/Qualifiers
 source 1..1323

```

/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"
1. .1323
/notes="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC21810.1"
/db_xref="GI:12226904"
/translation="MKAILIPFLSLIPLTPQSAFAQSEPELKLESVVIIVSRHGVRA
TKATQLMQDVTTPDAMPVTKLGLWTPRGELIAYLGHYQRLVADGLLAKGCPQS
GOVAIADVERKTRDTGEFAAGLAPDCAITVHTOADTSSPDPLEFNPLKTVGCOLDNA
NVTDAIIRAGGSTADFTGROTAFLERVLNFPQSNLCIKRKQDESCLTQALPS
ELKVSADNVSITGAVSLASMLTIFILQQAQMPGPGWRITDSHWNLTLSLHNAQF
YLQRTDEVARASPTPLLDLMAALTTPHPKQAYGVTLPTSLVLFIAQHDNLNANLGG
ALBNLWTLPOQDNTPPGGLVFRWRRLSDNSQWIQVSLVFTLQOMRDKTFLSLNT
PPGEVKLTLAGCBERNAGMCSLAGFTQIVNEARIIPACSLRSHHHHH"
ORIGIN
Alignment Scores:
Pred. No.:      8.92e-205      Length:      1323
Score:          2302.00        Matches:      440
Percent Similarity: 100.0%      Conservative:  0
Best Local Similarity: 100.0%      Mismatches:   0
Query Match:      100.0%      Indels:       0
DB:               2           Gaps:          0

US-10-601-319-2 (1-440) x AX052713 (1-1323)
Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAGCGATCTTAATCCCATTTTATCTCTTCGTGATTCGTTAACCCCGCATCTGCA 60
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db 61 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGAAGTGTGGTATTGTTCAGTCGTATGGT 120
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 121 GTGCGTCTCAACAAGGCCACGCAACTGATGAGGATGTACCCCGACGATGGCCA 180
Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
Db 181 ACCTGGCGGTAAACCTGGGTGGCTGACACCGCGGTGGTGAGCTAAATCGCCTATCTC 240
Qy 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 241 GGACATTTACCAAGCCAGCGTCTGTAGCCGACGAGTTGCTGGCGAAAGGGCTGCCCG 300
Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 301 CAGTCGTGTCAGGTCGCGGATTTATTCGTGATGTGCAGAGCGGTACCCGTAACAACAGGCGAA 360
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCCTTCGCGCGCGGCTGGACCTGTGTCGATACCGTACATACCGAGCGAGATACG 420
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCCAGTCCCGATCGTATTATTAATCCTCTAAACCTGGCGTTTCCCAACTGGATAACGCG 480
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGACTGACGCGATCCTCAGCAGGCGAGGAGGTCAATTGTGTGACTTTACCGGCGAT 540
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 541 CGGCAACCGGCTTTCGCGAACTGGAACGGGTGTTAATTTTCCGCAATCAAACTTGTGC 600
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCTATTACCGAGGCAATTTACCATCGGAATC 660
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"
1. .1323
/notes="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC21810.1"
/db_xref="GI:12226904"
/translation="MKAILIPFLSLIPLTPQSAFAQSEPELKLESVVIIVSRHGVRA
TKATQLMQDVTTPDAMPVTKLGLWTPRGELIAYLGHYQRLVADGLLAKGCPQS
GOVAIADVERKTRDTGEFAAGLAPDCAITVHTOADTSSPDPLEFNPLKTVGCOLDNA
NVTDAIIRAGGSTADFTGROTAFLERVLNFPQSNLCIKRKQDESCLTQALPS
ELKVSADNVSITGAVSLASMLTIFILQQAQMPGPGWRITDSHWNLTLSLHNAQF
YLQRTDEVARASPTPLLDLMAALTTPHPKQAYGVTLPTSLVLFIAQHDNLNANLGG
ALBNLWTLPOQDNTPPGGLVFRWRRLSDNSQWIQVSLVFTLQOMRDKTFLSLNT
PPGEVKLTLAGCBERNAGMCSLAGFTQIVNEARIIPACSLRSHHHHH"
ORIGIN
Alignment Scores:
Pred. No.:      8.92e-205      Length:      1323
Score:          2302.00        Matches:      440
Percent Similarity: 100.0%      Conservative:  0
Best Local Similarity: 100.0%      Mismatches:   0
Query Match:      100.0%      Indels:       0
DB:               2           Gaps:          0

US-10-601-319-2 (1-440) x AX052713 (1-1323)
Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAGCGATCTTAATCCCATTTTATCTCTTCGTGATTCGTTAACCCCGCATCTGCA 60
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db 61 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGAAGTGTGGTATTGTTCAGTCGTATGGT 120
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 121 GTGCGTCTCAACAAGGCCACGCAACTGATGAGGATGTACCCCGACGATGGCCA 180
Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
Db 181 ACCTGGCGGTAAACCTGGGTGGCTGACACCGCGGTGGTGAGCTAAATCGCCTATCTC 240
Qy 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 241 GGACATTTACCAAGCCAGCGTCTGTAGCCGACGAGTTGCTGGCGAAAGGGCTGCCCG 300
Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 301 CAGTCGTGTCAGGTCGCGGATTTATTCGTGATGTGCAGAGCGGTACCCGTAACAACAGGCGAA 360
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCCTTCGCGCGCGGCTGGACCTGTGTCGATACCGTACATACCGAGCGAGATACG 420
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCCAGTCCCGATCGTATTATTAATCCTCTAAACCTGGCGTTTCCCAACTGGATAACGCG 480
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGACTGACGCGATCCTCAGCAGGCGAGGAGGTCAATTGTGTGACTTTACCGGCGAT 540
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 541 CGGCAACCGGCTTTCGCGAACTGGAACGGGTGTTAATTTTCCGCAATCAAACTTGTGC 600
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCTATTACCGAGGCAATTTACCATCGGAATC 660

```

GOVATIDVDERTRKTBFAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQLDNA
NVTDAISRAGGSADFTGHRQTAFRELEQVNFQSNLCKRKRQSCSLTQALPS
ELKVSADNVSLTGAVSLAMLTETFLAQOQMPPEPGWGRITTDHSHWNTLLSHNAQF
YLLQRTPEVARSRAPIPLDLMAALTTPHPQKQAGVTLPTSLVPLFIAGHDNLANLGG
ALELNTWLTGPGPDNTPPGELVFERWRRLSDNSQWTVSLVFTLQQRDKTPIUSLNT
PFGEVKLTLAGEERNAQGMCSLAGFTQIVNEARIPACSLRSHHHHH"

ORIGIN

Alignment Scores:

Pred. No.: 8, 92e-205 Length: 1323
Score: 2302.00 Matches: 440
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-601-319-2 (1-440) x AX338966 (1-1323)

```
QY 1 MetLysAlaLeuLeuProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
DB 1 ATGAAGCGATCTTAATCCATTTTATCTCTCTGATTCCGTTAAACCCCGCAATCTGCA 60

QY 21 PheAlaGlnSerGluProGlnLeuLysLeuGluSerValValIleValSerArgHisGly 40
DB 61 TTCGCTCAGAGTGAGCCGAGCTGAAGCTGGAAAGTGTGGTATGTGTCAGTCTCATGCT 120

QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
DB 121 GTGCGTGTCTCCAAACAGCCACGCACTGATGCGAGGATGTACCCAGAGCGCATGGCCA 180

QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
DB 181 ACCTGCGCGTAAACACTGGTGTGCTGACCGCGNGGTGGTGAGTCAATCGCCTATCTC 240

QY 81 GlyHisThrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
DB 241 GGACATTACCAACGCGCAGCGTCTGTAGCCGACGATGTGCTGGCGAAGAGGCTGCCCG 300

QY 101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
DB 301 CAGTCTGTCAGTCCGATATTGCTGATGTGACGAGCGGTACCCGTAAACAGCGGAA 360

QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
DB 361 GCCITTCGCGCGCGGCTGCACCTGACTGTGCAATAACCGTACATACCAGCGCAGATAG 420

QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAla 160
DB 421 TCCAGTCCCGATCCGTTATTTAATCTCTAATAAAGCTGGCGTTTGCAACTGGATAACGCG 480

QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
DB 481 AACGTGACTGACCGCATCTCAGCAGGCGCAGAGGTCATTTGCTGACTTTACCGGGCAT 540

QY 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
DB 541 CGGCAACCGCGGTTTCGGAACCTGGAACCGGTGCTTAATTTTCGCAATCAACTGTGCG 600

QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
DB 601 CTTAAACGTGAGAAACAGCAGCAAGCTGTTCAATTAACGAGGCAATTAACCATCGGAATC 660

QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
DB 661 AAGGTGAGCGCGGACAAATGTCTCAATTAACCGGTGCGGTAAACCTCGCATCAATGCTGAG 720

QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
DB 721 GAGATATTTCTCTCGCAACAGCAGCAGGGAATGCCGAGCGCGGTGGGAGAGATACCC 780

QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
DB 781 GATTACACACGAGTGAACACCTTGCTAAGTTTGCATAACGCGCAATTTATTTGCTACAA 840
```

```
QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
DB 841 CGCAGCGCAGAGGTTGCCCGCAGCGGCCACCCCGTTATTGGATTGATCATGGCAGCG 900

QY 301 LeuThrProHisProProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
DB 901 TTGACGCCCATCCACCGCAAAAACAGCGGTATGGTGTGACATTACCCACTTCAGTACTG 960

QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp 340
DB 961 TTTATTGCGGACACGATACATAATCTGGCAAACTCCGCGCGCAGCTGGAGCTCAACTGG 1020

QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
DB 1021 ACGTTTCCCGGTGACCGCGATACACGCGCCAGTGGTGAACCTGGTGTGACGCTGG 1080

QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
DB 1081 CGTCGGCTAAGCGATAACAGCCAGTCAGATTGAGTTTCGCTGGTCTTCCAGACTTTACAG 1140

QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
DB 1141 CAGATGCGGTGATAAACCAGCCGCTGTCATTAAATACGCCCGCGGAGAGGTGAACCTGACC 1200

QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
DB 1201 CTGCGCAGATGTGAAGAGCGAATTCGCGAGGCAATGTTGTTGGCAGGTTTACGCA 1260

QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
DB 1261 ATCGTGAATGAAGCAGCATACCGCGTGCAGTTTGAGATCTCATCACCATCACCATCAC 1320

RESULT 8
AX356566 LOCUS 1323 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 1 from Patent WO0190333.
ACCESSION AX356566
VERSION AX356566.1 GI:18621053
KEYWORDS Escherichia coli
SOURCE Escherichia coli
ORGANISM Escherichia coli
REFERENCE 1 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
AUTHORS Enterobacteriaceae; Escherichia.
TITLE Short, J.M., Kretz, K.A., Gray, K.A., Barton, N.R., Garrett, J.B. and
JOURNAL O'Donoghue, E.
Recombinant bacterial phycases and uses thereof
Patent: WO 0190333-A 1 29-NOV-2001;
DIVERSA CORPORATION (US)
FEATURES
source Location/Qualifiers
1. 1323
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"
1. 1323
/notes="unnamed protein product"
CDS
1. 1323
/codon_start=1
/transl_table=11
/protein_id="CAD23025.1"
/db_xref="GI:18621054"
TRANSLATION="MKAILIPLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRA
P
TKATLQMDVTPDANPTWPKLWLPFGGELIAYLGHYORQLRVADGLLAKGCGPOS
GQVATIDVDERTRKTBFAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQLDNA
NVTDAISRAGGSADFTGHRQTAFRELEQVNFQSNLCKRKRQSCSLTQALPS
ELKVSADNVSLTGAVSLAMLTETFLAQOQMPPEPGWGRITTDHSHWNTLLSHNAQF
YLLQRTPEVARSRAPIPLDLMAALTTPHPQKQAGVTLPTSLVPLFIAGHDNLANLGG
ALELNTWLTGPGPDNTPPGELVFERWRRLSDNSQWTVSLVFTLQQRDKTPIUSLNT
PFGEVKLTLAGEERNAQGMCSLAGFTQIVNEARIPACSLRSHHHHH"
ORIGIN
Alignment Scores: 8, 92e-205 Length: 1323
Pred. No.: 8, 92e-205 Length: 1323
```


Db 428 GGACATTACCAACGCCAGCGTCTGGTAGCCGCGATTGCTGGCGAAAGGCGTCCCG 487
Qy 101 GlnSerGlyGlnValAlaIleAlaAaspValAaspGluAargThrArgLysThrGlyGlu 120
Db 488 CAGTCTGTGTCAGTGGCGATTATGCTGATGTCGACGAGCGTACCCTGAAACAGCGGAA 547
Qy 121 AlaPheAlaAlaGlyLeuAlaProAaspCysAlaIleThrValHisThrGlnAlaAaspThr 140
Db 548 GCCTTCGGCGCGCGTGGCCTGACTGTGTCGCAATACCGTACATACCCAGCCAGATACG 607
Qy 141 SerSerProAaspProLeuPheAaspProLeuLysThrGlyValCysGlnLeuAaspAla 160
Db 608 TCCAGTCCGAGTCCGCTTATTAATCTCTAAAAACCTGGCGTTGCAACTGGATACGCG 667
Qy 161 AsnValThrAaspAlaIleLeuSerArgAlaGlyGlySerIleAlaAaspPheThrGlyHis 180
Db 668 AACGTGACTGACCGCATCTCAGCAGGCGAGAGGGTCAATTCCTGACTTTACCGGGCAT 727
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAaspPheProGlnSerAsnLeuCys 200
Db 728 CGGCAAAACGGCGTTTCGGGAACTGGAAACGGGTGCTTAATTTTCGCGCAATCAACTTGTGC 787
Qy 201 LeuLysArgGluLysGlnAaspGlnSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 788 CTTAAACCTGAGAAACAGGACGAAAGCTGTTCAATTAACGCGACGATTAACCATCGGAATC 847
Qy 221 LysValSerAlaAaspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 848 AAGGTGAGCGCGCAATGCTTCATTAACCGGTGCGTAAGCTCGCATCAATGCTGACG 907
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
Db 908 GAGATATTTCTCTCAACACAGCAGGGAATGCCGAGCGCGGGTGGGGAAGGATCACC 967
Qy 261 AspSerHisGlnThrAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 968 GATTCACACCAAGTGAACACCTTGTCAAGTTTGCAATAACGCGCAATTTATTTGTCTACA 1027
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAaspLeuIleMetAlaAla 300
Db 1028 CGCAGCCAGAGGTGGCCGAGCGCGCCACCGCTTATTAGATTGATCAAGACAGCG 1087
Qy 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 1088 TTGACGCCCATCCACCGCAAAACAGCGGTATGTTGTGACATTAACCACTTCAGTGTG 1147
Qy 321 PheIleAlaGlyHisAaspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAaspThr 340
Db 1148 TTTATCGCGGACACGATACTAATCTGGCAAACTCTCGCGCGCGCACTGGAGCTCAACTGG 1207
Qy 341 ThrLeuProGlyGlnProAaspAsnThrProProGlyGlyGluLeuValPheGluArgTyr 360
Db 1208 ACGCTTCCTCGGTACGCGGATTAACACGCGCGAGGTGTGAACCTGTTTGAACGCTGG 1267
Qy 361 ArgArgLeuSerAaspAsnSerGlnTyrIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1268 CGTGGCTAAGCGATAACAGCAGCGATGATTCAGTTTCGCTGGTCTTCCAGCTTTACG 1327
Qy 381 GlnMetArgAaspLysThrProLeuSerLeuAaspThrProProGlyGluValLysLeuThr 400
Db 1328 CAGATGCGTGATAAAACCGCGCTGTCAATTAATAACGCGCGCGAGAGGTGAACACTGACC 1387
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1388 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTCTGTTGGCAGGTTTACGCA 1447
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1448 ATCGTGATGAAGCAGCATACCGCGGTGCGAGTTG 1483

RESULT 10
AX356572
LOCUS

AX356572 1901 bp DNA linear PAT 06-FEB-2002

DEFINITION Sequence 7 from Patent WO0190333.
ACCESSION AX356572
VERSION AX356572.1 GI:18621059
KEYWORDS
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE Short,J.M., Kretz,K.A., Gray,K.A., Barton,N.R., Garrett,J.B. and
O'Donoghue,E.
AUTHORS Recombinant bacterial phytases and uses thereof
TITLE Patent: WO 0190333-A 7 29-NOV-2001;
JOURNAL DIVERSA CORPORATION (US)
FEATURES Location/Qualifiers
source
1..1901
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"

ORIGIN

Alignment Scores:
Pred. No.: 2,74e-198 Length: 1901
Score: 2235.00 Matches: 430
Percent Similarity: 99.5% Conservative: 0
Best Local Similarity: 99.5% Mismatches: 2
Query Match: 97.1% Indels: 0
DB: 2 Gaps: 0

US-10-601-319-2 (1-440) x AX356572 (1-1901)

Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 188 ATGAAGCGATCTTAATCCCATTTTATCTCTTCGATTCCGTTAACCCCGAATCTGCA 247
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Db 248 TTCGCTCAGAGTAGCGGAGCTGAGCTGGAAGTGTGGTGTGTTGTCAGTCGTCATGGT 307
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAaspValThrProAaspAlaTyrPro 60
Db 308 GTGCGTGTCTCAACCAAGGCGCACCACTGATGAGGATGTCAACCCAGAGCGCATGGCA 367
Qy 61 ThrTrpProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db 368 ACCTGGCCGGTAAACCTGGGTGGCTGACACCGCGNGGTGGTGAAGTAAATCGCTTATCTC 427
Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAaspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 428 GGACATTACCAACGCCAGCGTCTGTAGCCGAGTGTGTCGGCGAAAGGCGTCCCG 487
Qy 101 GlnSerGlyGlnValAlaIleAlaAaspValAaspGluArgThrArgLysThrGlyGlu 120
Db 488 CAGTCTGCTGAGTCCGCGATTATTGCTGATGTGACGAGCGTACCCGTAACAGCGGAA 547
Qy 121 AlaPheAlaAlaGlyLeuAlaProAaspCysAlaIleThrValHisThrGlnAlaAaspThr 140
Db 548 GCCTTCGGCGCGCGCTGGCACCTGACTGTGCAATAACCGGTGCTTCCCAACTGGATAACGCG 607
Qy 141 SerSerProAaspProLeuPheAaspProLeuLysThrGlyValCysGlnLeuAaspAla 160
Db 608 TCCAGTCCCGCATCGTTATTAATCTCTAAAAACCTGGCGTTGCGCAACTGGATAACGCG 667
Qy 161 AsnValThrAaspAlaIleLeuSerArgAlaGlyGlySerIleAlaAaspPheThrGlyHis 180
Db 668 AACGTGACTGACCGCATCTCAGCAGGCGAGGAGGTCAATTCCTGACTTTTACCGGGCAT 727
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAaspPheProGlnSerAsnLeuCys 200
Db 728 CGGCAAAACGGCGTTTCGCGAACTGGAAACGGGTGCTTAATTTTCGCGCAATCAACTTGTGC 787
Qy 201 LeuLysArgGluLysGlnAaspGlnSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220

Db	788	CTTAAACGTGAGAAACAGGACGAAAGCTGTTTCATTAACGAGGCATTACCATCGGAATC	847
Qy	221	LyvValSerAlaAapAnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr	240
Db	848	AAGGTGAGCCGCGACATGCTCTATTAAACGGTGCCTAAGCCTCGCATCAATGCTGACG	907
Qy	241	GlulilePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyVArgIleThr	260
Db	908	GAGATATTTCTCTGCAACAGCACAGGAAATGCCGAGCGGGGTGGGAGAGATCAC	967
Qy	261	AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln	280
Db	968	GATTACACACAGTGGACACCTTCTAAGTTTGATTAACGCGCAATTTTATTTTGCTACAA	1027
Qy	281	ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla	300
Db	1028	CGCACGCGCAGAGGTGGCCGAGCGCGCCACCCCGTTATTAGATTTTGATCAAGACAGCG	1087
Qy	301	LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu	320
Db	1088	TTGACGCCCCATCCACCGCAAAACAGCGGTATGCTGATACATTTACCCACTTCAGTGCTG	1147
Qy	321	PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp	340
Db	1148	TTTATCGCGGACACGATACTAATCTGGCAATCTCGCGCGCGCACTGGAGCTCAACTGG	1207
Qy	341	ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp	360
Db	1208	ACGCTTCGCGGTACGCGGATTAACACGCGCGAGTGGTGAACCTGGTGTGTAACGCTGG	1267
Qy	361	ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln	380
Db	1268	CGTCGGCTAAGCGATAAACAGCCAGTGGATTGAGTTTCGGTGTTCACGACTTTACAG	1327
Qy	381	GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr	400
Db	1328	CAGATGCGTGATAAACCGCGCTGCTATTAATAACGCGCGCGAGAGGTGAACCTGACC	1387
Qy	401	LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln	420
Db	1388	CTGCGAGGATGTGAAGCGCAAAATCGCGAGGCATGTGTGTTGGCAGGTTTTTACGCA	1447
Qy	421	IleValenGluAlaArgIleProAlaCysSerLeu	432
Db	1448	ATCGTGATGAAGCACGCATACCGCGGTGCAGTTTG	1483
RESULT 11			
ECOAPPA		1901 bp	DNA linear
LOCUS			BCT 04-APR-2002
DEFINITION			Escherichia coli periplasmic phosphoanhydride phosphohydrolase
ACCESSION			(appA) gene, complete cds.
VERSION			M58708
KEYWORDS			M58708.1 GI:145283
SOURCE			Escherichia coli
ORGANISM			Escherichia coli
REFERENCE			Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
AUTHORS			Enterobacteriaceae; Escherichia.
TITLE			1 (bases 1 to 1901)
			Daasa,J., Marck,C. and Boquet,P.L.
			The complete nucleotide sequence of the Escherichia coli gene appA
			reveals significant homology between pH 2.5 acid phosphatase and
			Glucose-1-phosphatase
JOURNAL			J. Bacteriol. 172 (9), 5497-5500 (1990)
PUBMED			2168385
REFERENCE			2 (sites)
AUTHORS			Ostanin,K., Harms,E.H., Stevis,P.E., Kuciel,R., Zhou,M.M. and Van
			Eten,R.L.
			Overexpression, site-directed mutagenesis, and mechanism of
			Escherichia coli acid phosphatase
JOURNAL			J. Biol. Chem. 267 (32), 22830-22836 (1992)
PUBMED			1429631
REFERENCE			3 (sites)

AUTHORS	Ostanin,K. and Van Etten,R.L.
TITLE	Asp304 of Escherichia coli acid phosphatase is involved in leaving group protonation
JOURNAL	J. Biol. Chem. 268 (28), 20778-20784 (1993)
PUBMED	8407904
FEATURES	
source	Location/Qualifiers
	1..1901
	/organism="Escherichia coli"
	/mol_type="genomic DNA"
	/db_xref="taxon:562"
gene	3..1522
	/gene="appA"
RBS	3..8
	/gene="appA"
CDS	16..108
	/gene="appA"
	/note="ORFX"
	/codon_start=1
	/transl_table=11
	/product="unknown"
	/protein_id="AAA72085.1"
	/db_xref="GI:145284"
	/translation="MWYLLWFVGILLMCSLSLTLVLVWLDPRLLKS"
stem_loop	118..149
	/gene="appA"
-35_signal	133..138
	/gene="appA"
-10_signal	156..161
	/gene="appA"
RBS	172..177
	/gene="appA"
CDS	188..1486
	/gene="appA"
	/standard_name="acid phosphatase"
	/codon_start=1
	/transl_table=11
	/product="periplasmic phosphoanhydride phosphohydrolase"
	/protein_id="AAA72086.1"
	/db_xref="GI:145285"
	/translation="WKAILIPFLSLIPLTPOSAPFAQSEPELKESVIVSRHGVRA TKATQIMQDVTPDAMTWPVKLGMLTTPRGELIAYLGHYQROLVADGLAKKGCPS GQVAIIADVDERTKTRTGEAPAGLAPDCAITVHTQADTSPPDLPFLPKTGVQQLDNA NVTDAILSRAGSGIADFTGHROTAFRELERNLFPQSNCLKREKODECSLTOALPS EKLVSADNVSLTCAVSLASMLTEIFLLQAOQMPGPWGRIITDSHOWNTLLSHNAQF YLLQRTPEVARSRATPELLDLIKTALTPHPQKQAYGVLTPTSVLFTAGHDTNLNLGG ALELNLPLGCPDPTPPGGELVPERWRRLSDNSQWIOVSLVFTLQOMRDKTFLSLNT PPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIIPACSL"
sig_peptide	188..253
	/gene="appA"
mat_peptide	254..1483
	/gene="appA"
	/product="periplasmic phosphoanhydride phosphohydrolase"
variation	299..301
	/gene="appA"
	/standard_name="H303A mutant"
	/note="created by site directed mutagenesis"
	/citation=[2]
variation	299..301
	/gene="appA"
	/standard_name="R16A mutant"
	/note="created by site directed mutagenesis"
	/citation=[2]
variation	299..301
	/gene="appA"
	/standard_name="R20A mutant"
	/note="created by site directed mutagenesis"
	/citation=[2]
variation	299..301
	/gene="appA"
	/standard_name="R92A mutant"
	/note="created by site directed mutagenesis"
	/citation=[2]
variation	302..304

```

/gene="appa"
/standard_name="H17N mutant"
/note="created by site directed mutagenesis"
/citation={2}
1163. .1165
/gene="appa"
/standard_name="D304A mutant"
/note="created by site directed mutagenesis"
/citation={3}
/phenotype="Asp 304 changed to Ala"
/replace="gct"
1163. .1165
/gene="appa"
/standard_name="D304E mutant"
/note="created by site directed mutagenesis"
/citation={3}
/phenotype="Asp 304 changed to Glu"
/replace="gag"
1497. .1522
/gene="appa"
/note="rho-independent terminator"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 2.74e-198 Length: 1901
Score: 2235.00 Matches: 430
Percent Similarity: 99.5% Conservative: 0
Best Local Similarity: 99.5% Mismatches: 2
Query Match: 97.1% Indels: 0
DB: 15 Gaps: 0

```

US-10-601-319-2 (1-440) x ECOAPPAA (1-1901)

```

Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
Db 198 ATGAAGCGAICTTAATCCATTTTATCTCTCTGATTCCGTTAAACCCCGCAATCTGCA 247
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40
Db 248 TTGCTCAGAGTGAGCCGAGCTGAAGCTGAAGAGTGTTGATTTGTCAGTCGTCATGTT 307
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 308 GTGCGTGTCTCCAAACCAAGCCACGCACTGATGCAGGATGTCAACCCAGAGCGCATGGCCA 367
Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrlleu 80
Db 368 ACCTGCGCCGGTAAAACTGGGTTGGCTGACACCGGNGGTGGTGAGCTAATCGCCTATCTC 427
Qy 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 428 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGGATTGCTGGCGAANAAGGGCTGCCCG 487
Qy 101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 488 CAGTCTGCTCAGTGGCGGATTAATGCTGATGTCGACGAGCGTACCCGTAAACAGGCGAA 547
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCCTTCGCGCGCGGCTGGACCTGACTGTGCAATAACCGTACATACCCAGCGAGATACG 607
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 608 TCCAGTCCCGATCCGTTAATTAATCTCTAAAACTGGCGTTGGCAACTGGATAACGCG 667
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db 668 AACGTGACTGACGGATCCTCAGCAGGCGCAGGAGGTCAATTGCTGACTTTACCGGGCAT 727
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 728 CGGAAACGGCGGTTTCGGCAACTGGACGGGTGCTTAATTTTCCGCAATCAAACTTGTC 787

```

```

Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 788 CTTRAAACGTGAGAAACAGGACGAAAGCTGTTCTATTAAACGAGGCATTACCATCGGAATC 847
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 848 AAGGTGAGCGCGCAATGTCTCTATTAAACCGGTGCGGTGAAGCTCGCATCATGTGTGAG 907
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Db 908 GAGATATTCTTCCTGCAACAAGCACAGGGAATGCCGAGCGCGGGTGGGGAAGGATCAC 967
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
Db 968 GATTACACACAGTGAACACACTTGTCTAAGTTTGATTAACGCGCAATTTTATTGTCTACAA 1027
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 1028 CGCACGCGCAGAGTTGCCCGCAGCGCCGCCACCCCGTTATTAGATTTCATCAAGACAGCG 1087
Qy 301 LeuThrProHisProGlnLysGlnAlaTyrlGlyValThrLeuProThrSerValIle 320
Db 1088 TTGACGCCCCATCCACCAAAACAGGCGTATGTGTGACATTACCCACTTCAGTGTGTG 1147
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 1148 TTTATCGCGGACAGCATACTTAATCTGGCAATCTCGCGGGGCACTCGAGCTCAACTGG 1207
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 1208 ACGCTTCGCGTCAGCGGATAACACGCGCAGGTGTTGTAACCTGTTTGAACGCTGG 1267
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1268 CGTGGCTAGCGGATAACAGCCAGTGGATTGAGTTTCGCTGGTTCCTCCAGCTTTACAG 1327
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 1328 CAGATGCTGATAAAACGCGCTGTCATTAATAACGCGCCCGGAGAGGTGAAACTGACC 1387
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1388 CTGCGCAGGATGTGAAGAGCGAAATGCGCAGGCGATGTTGTTGTTGGCAGGTTTACGCAA 1447
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1448 ATCGTGAATGAAGCACGCATACCGCGCTGCAGTTTG 1483

RESULT 12
AX042376
LOCUS AX042376 3470 bp DNA linear PAT 23-NOV-2000
DEFINITION Sequence 5 from Patent WO0064247.
ACCESSION AX042376
VERSION AX042376.1 GI:11340994
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Forsberg, C.W., Golovan, S. and Phillips, J.P.
TITLE Transgenic animals expressing salivary proteins
JOURNAL Patent: WO 0064247-A 5 02-NOV-2000;
UNIVERSITY OF GUELPH (CA)
FEATURES
    source
        1. .3470
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="R15/APPA transgene"
ORIGIN
Alignment Scores: 6.59e-198 Length: 3470
Pred. No.:

```

```
Score: 2235.00 Matches: 430
Percent Similarity: 99.5% Conservative: 0
Best Local Similarity: 99.5% Mismatches: 2
Query Match: 97.1% Indels: 0
DB: 2 Gaps: 0

US-10-601-319-2 (1-440) x AX042376 (1-3470)

Qy 1 MetLysAlaLeuLeuLeuProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
Db 1811 ATGAAGCCATCTTAATCCCAATTTTATCTCTGATTCGGTTAAACCCGCAATCTGCA 1870
Qy 21 PheAlaGlnSerGluProGluLeuLeuLeuSerValValLeuValSerArgHisGly 40
Db 1871 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGGTATGTTCAGTCGTCATGGT 1930
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTyrPro 60
Db 1931 GTGGTGCTCCACCAAGGCCACCACTGATGACAGATGTACCCAGACGCATGGCCA 1990
Qy 61 ThrTrpProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuLeuAlaTyrLeu 80
Db 1991 ACCTGGCGGTAAACCTGGGTGGCTGACACCGCGGTGGTGAGCTTAATCGCCTATCTC 2050
Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 2051 GGACATTAACCAACCCAGCGCTCGGTAGCCGACGGATGTGGCGAAGGCGTGGCCG 2110
Qy 101 GlnSerGlyGlnValAlaLeuLeuAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 2111 CAGTCTGGTCAGTTCGGATTTATGCTGTGATGTGACGAGCGTACCCGTTAAACAGGCGAA 2170
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaLeuThrValHisThrGlnAlaAspThr 140
Db 2171 GCCTTCGCGCGCGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 2230
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 2231 TCCAGTCCCGATCCGTTATTAAATCCCTCTAAAACTCGCGTTTCCCAACTGGATTAACGCG 2290
Qy 161 AsnValThrAspAlaLeuSerArgAlaGlySerLeuAlaAspPheThrGlyHis 180
Db 2291 AACGTGACTAGCGATCTCTCAGCAGGCGAGGAGGTCAATGTCTGACTTTACCGGCAT 2350
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 2351 CGGCAACCGCGTTTCGCAACTGCAACCGGTGCTTAATTTTCGCAATCAAACTTGTC 2410
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 2411 CTTAAACGTGAGAAACACAGGACGAAAGCTGTTCAATTAACGAGGCATTAACCATCGAACTC 2470
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 2471 AAGGTGAGCGCGCAATGCTCTAATACCGTGGGTAGCCCTCGCATCATGCTGACG 2530
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
Db 2531 GAGATATTTCTCTGCAACAGCACAGGAATGCGGAGCGCGGTGGGAAGGATCAAC 2590
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
Db 2591 GATTCAACACCATGTCACCTTCCTAAGTTTGCATTAACGCGCAATTTTATTTGCTACAA 2650
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuLeuMetAlaAla 300
Db 2651 CGACAGCGAGGTTGCGCGAGCGCGCCACCCCGCTTATAGATTTGATCAACAGACGCG 2710
Qy 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 2711 TTGACGCGCCCATCCACGCAAAAAACAGCGGTATGCTGTGACATTAACCACTTCAGTCTG 2770
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuLeuAsnTrp 340
```

```
Db 2771 TTTATCGCGGACACGATACTAATCTGGCAAAATCTCGCGCGCAGCTGGAGCTCAACTGG 2830
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTyr 360
Db 2831 ACGTTCCCGGTACGCGGATACACGCGCAGGTGGTGAACCTGGTGTGAAACGCTGG 2890
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 2891 CGTCGGCTAAGCGATAACACGCCAGTGGATTCAGGTTTCGCTGCTTCCAGACTTTACAG 2950
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 2951 CAGATGCGTGATAAAAACGCGCTGTCTAATAATACGCGCCCGAGAGGTGAACACTGACC 3010
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 3011 CTGGCAGGATGTGAAGCGAATATGCGAGGCGATGTTCGTTGGCAGGTTTACGCAA 3070
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 3071 ATCGTGAATGAAGCAGCATACCGCTTGCAAGTTTG 3106

RESULT 13
AX042374 LOCUS AX042374 4060 bp DNA linear PAT 23-NOV-2000
DEFINITION Sequence 3 from Patent WO0064247.
ACCESSION AX042374
VERSION AX042374.1 GI:11340992
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Forsberg, C.W., Golovan, S. and Phillips, J.P.
TITLE Transgenic animals expressing salivary proteins
JOURNAL Patent: WO 0064247-A 3 02-NOV-2000;
UNIVERSITY OF GUELPH (CA)
FEATURES
Location/Qualifiers
source 1. 4060
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="R15/APPA + intron transgene"

ORIGIN

Alignment Scores:
Pred. No.: 8 28e-198 Length: 4060
Score: 2235.00 Matches: 430
Percent Similarity: 99.5% Conservative: 0
Best Local Similarity: 99.5% Mismatches: 2
Query Match: 97.1% Indels: 0
DB: 2 Gaps: 0

US-10-601-319-2 (1-440) x AX042374 (1-4060)

Qy 1 MetLysAlaLeuLeuLeuProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
Db 1811 ATGAAGCCATCTTAATCCCAATTTTATCTCTGATTCGGTTAAACCCGCAATCTGCA 1870
Qy 21 PheAlaGlnSerGluProGluLeuLeuLeuSerValValLeuValSerArgHisGly 40
Db 1871 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGGTATGTTCAGTCGTCATGGT 1930
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTyrPro 60
Db 1931 GTGGTGCTCCACCAAGGCCACCACTGATGACAGATGTACCCAGACGCATGGCCA 1990
Qy 61 ThrTrpProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuLeuAlaTyrLeu 80
Db 1991 ACCTGGCGGTAAACCTGGGTGGCTGACACCGCGGTGGTGAGCTTAATCGCCTATCTC 2050
Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
```

```
Db 2051 GGACATTACCAACGCGCGTCTGGTAGCCGCGATTGTCGGCAAAAGGGCTGCCCG 2110
Qy 101 GlnSerGlyGlnValAlaIleAlaAaspValAspGluAtrGthrArgLysThrGlyGlu 120
Db 2111 CAGTCTGGTCAGTCCGATTAATGCTGATGTCGACGAGCGTACCCTGAAACAGCGCAA 2170
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAaspThr 140
Db 2171 GCCTTCGCGCGCGCTGGCACTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 2230
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 2231 TCCAGTCCCGATCCGCTTAATTAATCCTCTAAACACTGGCGTTGTCACATGGATAACGCG 2290
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAaspPheThrGlyHis 180
Db 2291 AACGTGATGACGCGATCTCAGCAGGCGGAGAGGGTCAATGCTGATTTACCGGGCAT 2350
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 2351 CGGCAAAACGGCGTTTCGGCAACTGGAAACGGGTGCTTAATTTTCGGCAATCAACTTGTGC 2410
Qy 201 LeuLysArgGluLysGlnAaspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 2411 CTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGACGAGCATTAACCATCGGAATC 2470
Qy 221 LysValSerAlaAaspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 2471 AAGGTGACGCGCGACATGCTCTAATAACCGGTGGGTGAAGCTCGCATCAATGCTGACG 2530
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnMetProGluProGlyLysThrGlyValArgIleThr 260
Db 2531 GAGATATTTCTCTGCAACAGCAGGAGATGCCGAGCGCGGTGGGAGGATCAACC 2590
Qy 261 AspSerHisGlnThrAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 2591 GATTACACACAGTGGAAACACCTTGCTAAGTTTGCAATAACGGCGCAATTTATTTGTACAA 2650
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 2651 CGCAGCCAGAGGTTGCCCGCAGCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCG 2710
Qy 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 2711 TTGACGCCCCATCCACCGCAAAACAGCGGTATGTTGTGACATTAACCCACTTCAGTGTG 2770
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 2771 TTTATCGCGGACACAGATACTAATCTGGCAAACTCTCGCGCGCACCTGGAGCTCAACTGG 2830
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 2831 ACGCTTCGCGTACGCGGATACACGCGCGCGAGGTGGTGAACCTGGTGTGTTGAACGCTGG 2890
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 2891 CGTGGCTAAGCGATTAACGACGAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTTACG 2950
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 2951 CAGATGCGTGATAAAACCGCGCTGTCTAATAACGCGCGCGAGAGGTGAACCTGACC 3010
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnMetCysSerLeuAlaGlyPheThrGln 420
Db 3011 CTGGCAGATGTGAAGAGCGAAATCGCAGGCGCATGTGTTGTTGGCAGGTTTTCACGCAA 3070
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 3071 ATCGTGAATGAAGCAGCATACCCCTTCGAGTTTG 3106
```

RESULT 14
AX042375

```
LOCUS AX042375 6116 bp DNA linear PAT 23-NOV-2000
DEFINITION Sequence 4 from Patent WO0064247.
ACCESSION AX042375
VERSION AX042375.1 GI:11340993
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Forsberg C.W., Golovan,S. and Phillips,J.P.
TITLE Transgenic animals expressing salivary proteins
JOURNAL Patent: WO 0064247-A 4 02-NOV-2000;
UNIVERSITY OF GUELPH (CA)
FEATURES
source
1. 6116
Location/Qualifiers
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="R15/APPA plasmid with pBLCAT3 vector"
```

ORIGIN

Alignment Scores:

Pred. No.:	1.51e-197	Length:	6116
Score:	2235.00	Matches:	430
Percent Similarity:	99.5%	Conservative:	0
Best Local Similarity:	99.5%	Mismatches:	2
Query Match:	97.1%	Indels:	0
DB:	2	Gaps:	0

US-10-601-319-2 (1-440) x AX042375 (1-6116)

Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20

Db 1811 ATGAAGCCATCTTAATCCCATTTTATCTCTTCTGATTCCTTAACCCCGAATCTGCA 1870

Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40

Db 1871 TTCGCTCAGAGTAGCGGAGCTCAAGCTGAAGCTGGAAGTGTGGTGATTGTCAGTCGTCATGGT 1930

Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60

Db 1931 GTGGCTGCTCTCAACCAAGGCCACGCAACTGATGACAGGATGTCAACCCAGACGCAATGGCA 1990

Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80

Db 1991 ACCTGGCGGTAAACCTGGGTGGCTGACACCGCGCGGTGGTGAGCTAATCGCTAATCTC 2050

Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAaspGlyLeuLeuAlaLysLysGlyCysPro 100

Db 2051 GGACATTACCAACGCCAGCGTCTGGTAGCCGCGATTGCTGGCGAAAGGGCTGCCCG 2110

Qy 101 GlnSerGlyGlnValAlaIleIleAlaAaspValAspGluArgThrArgLysThrGlyGlu 120

Db 2111 CAGTCTGGTCAGTCCGCGATTAATGCTGATGTCGACGAGCGTACCCTGTAACAGCGGCA 2170

Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAaspThr 140

Db 2171 GCCTTCGCGCGCGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATACG 2230

Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160

Db 2231 TCCAGTCCCGATCCGTTAATTAATCCTCTAAACACTGGCGTTTCCCAACTGGATAACGCG 2290

Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAaspPheThrGlyHis 180

Db 2291 AACGTGATGACGCGATCTCAGCAGGCGGAGAGGGTCAATGCTGATTTACCGGGCAT 2350

Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200

Db 2351 CGGCAAAACGGCGTTTCGGCAACTGGAAACGGGTGCTTAATTTTCGGCAATCAACTTGTGC 2410

Qy 201 LeuLysArgGluLysGlnAaspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220

Db 2411 CTTAAACGTGAGAACAGGACGAAGCTGTTTCATTAAACGAGGCAATTACCATCGGAATC 2470
Qy 221 LysValSerAlaAspAenValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 2471 AAGGTGAGCGCGCAATGTCTCATTAACCGTCCGCTCGCATCAATGCTGACG 2530
Qy 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTyrGlyArgGileThr 260
Db 2531 GAGATATTCTCTCGCAACAGCACAGGAATCCGAGCGCGGGTGGGAAGATCACC 2590
Qy 261 AspSerHisGlnTrpAenThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 2591 GATTACACCACTGGAACACCTTGCTAAGTTTGATTAACGGCAATTTATTGCTACAA 2650
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 2651 CGCACGCGCAGAGGTGGCCGAGCGCGCCACCCCGTTATTAGATTGTGATCAAGACAGCG 2710
Qy 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 2711 TTGACGCCCATCCACCGCAAAACAGCGGTATGGTGACATTTACCACCTTCAGTGCTG 2770
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 2771 TTTATCGCGGACACGATACTAATCTGGCAATCTCGCGCGCGCACTGGAGCTCAACTGG 2830
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 2831 ACGETTCGCGTCAAGCGGATAACACGCGCCAGTGGTGAACCTGGTTTGAACGCTGG 2890
Qy 361 ArgArgLeuSerAspAenSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 2891 CGTCGGCTAAGCGGATAACAGCCAGTGGATTACAGTTTCGTGGTCTTCACAGCTTTACAG 2950
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 2951 CAGATGGTGATAAAACCGCGCTGTCTAATAATAGCGCGCCCGAGAGGTGAACTGACC 3010
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 3011 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGCGATGTGTCTGTGGCAGGTTTTCACGAA 3070
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 3071 ATCGTGAATGAAGCACGATACCCGCTTCAGATTG 3106

RESULT 15
AX042373
LOCUS AX042373 6708 bp DNA linear PAT 23-NOV-2000
DEFINITION Sequence 2 from Patent WO0064247.
ACCESSION AX042373
VERSION AX042373.1 GI:11340991
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Forsberg,C.W., Golovan,S. and Phillips,J.P.
TITLE Transgenic animals expressing salivary proteins
JOURNAL Patent: WO 0064247-A 2 02-NOV-2000;
UNIVERSITY OF GUELPH (CA)
FEATURES
source location/Qualifiers
1..6708 /organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="R15/APPA + intron plasmid with pBLCAT3 vector"

ORIGIN

Alignment Scores:
Pred. No.: 1.72e-197 Length: 6708
Score: 2235.00 Matches: 430
Percent Similarity: 99.5% Conservative: 0

Best Local Similarity: 99.5% Mismatches: 2
Query Match: 97.1% Indels: 0
DB: 2 Gaps: 0
US-10-601-319-2 (1-440) x AX042373 (1-6708)
Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
Db 1811 ATGAAGCCATCTTAATCCCATTTTATCTCTTCGATTCCGTTAAACCCGCAATCTGCA 1870
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValVallIleValSerArgHisGly 40
Db 1871 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGGTATTGTCACTCGTCATGGT 1930
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 1931 GTGCGTGTCTCAACCAAGCCACGCAACTGATGCAGGATGTACCCGACGATGCGCA 1990
Qy 61 ThrTrpProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db 1991 ACCTGGCCGGTAAACCTGGGTTGCGCTGACACCGCGGGTGGTGAAGTAAATCGCCTATCTC 2050
Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 2051 GGACATTACCAACGCCAGCGTCTGGTAGCGAGGATTGCTGGGGAAGAGGGCTGCCG 2110
Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 2111 CAGTCTGGTTCAGGTCGCGATTATTGCTGATGTCCAGGCGTACCCGTAACAGGCGAA 2170
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 2171 GCCTTCGCGCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGAGATACG 2230
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 2231 TCCAGTCCCGATCCGTTATTATCTCTAANAACCTGGCGTTGCCAATGAGATACGCG 2290
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db 2291 AACGTGACTGACCGCATCTCTCAGCAGGCGAGGAGGTCAATTGCTGACTTTTACCGGCAT 2350
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 2351 CGGCAACAGCGCTTCGCGAACCTGGAACCGGTCTTAATTTCCGCAATCAAACTTGTGC 2410
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 2411 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGAGGCAATTACCATCGGAATC 2470
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 2471 AAGGTGAGCGCGCAATGTCTCATTAACCGGTGCGGTAAAGCTTCGATCAATGCTGACG 2530
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgGileThr 260
Db 2531 GAGATATTCTCTCGCAACAGCACAGGGAATCCGAGCGCGGGTGGGAAGATCACC 2590
Qy 261 AspSerHisGlnTrpAenThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 2591 GATTACACCACTGGAACACCTTGCTAAGTTTGATTAACCGCAATTTATTGCTACAA 2650
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 2651 CGCACGCGCAGAGGTGGCCGAGCGCGCCACCCCGTTATTAGATTGTGATCAAGACAGCG 2710
Qy 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 2711 TTGACGCCCATCCACCGCAAAACAGGCGGTATGGTGTGACATTACCCACTTCAGTGCTG 2770
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 2771 TTTATCGCGGACACGATACTAATCTGGCAATCTCGCGCGCGCACTGGAGCTCAACTGG 2830

QY	341	ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp	360
Db	2831	ACGCTTCCCGGTACGCCGATAACACACGCCGCGAGGTGGTGAACCTGGTGTGAAACGCTGG	2890
QY	361	ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln	380
Db	2891	CGTCGGCTAAGCGATAACAGCCAGTGGATTGAGGTTTCGCTGGTCTCCAGACTTTACAG	2950
QY	381	GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr	400
Db	2951	CAGATGCGTGATAAAACGCCGCTGTCAATTAATACGCCGCCGAGAGGTGAACCTGACC	3010
QY	401	LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln	420
Db	3011	CTGGCAGGATGTGAAGAGCGAAATGCCGAGGCGCATGTGTCTGTTGGCAGGTTTTACGCAA	3070
QY	421	IleValAsnGluAlaArgIleProAlaCysSerLeu	432
Db	3071	ATCGTGAATGAAGCACGCATACCCGCTTGCAGTTTG	3106

Search completed: June 14, 2006, 13:53:05
Job time : 5289.78 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2006, 10:22:24 ; Search time 25.7339 Seconds
(without alignments)
1645.118 Million cell updates/sec

Title: US-10-601-319-2
Perfect score: 2302
Sequence: 1 MKAILPFLSLILPLTPQSA.....IVNEARIPACSLRSHHHHH 440
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2235	97.1	432	2 B36733	acid phosphatase (
2	2210	96.0	434	2 H90770	phosphoanhydride p
3	2210	96.0	444	2 D85633	hypothetical prote
4	943.5	41.0	441	2 AC0201	acid phosphatase (
5	545	23.7	413	2 F90773	periplasmic glucos
6	545	23.7	413	2 B85636	periplasmic glucos
7	541	23.5	417	2 S25627	glucose-1-phosphat
8	539	23.4	413	2 JVQ087	glucose-1-phosphat
9	526	22.8	414	2 E87316	periplasmic phosph
10	524	22.8	413	2 AG0632	glucose-1-phosphat
11	139.5	6.1	423	1 S06167	acid phosphatase (
12	133	5.8	416	2 T16058	hypothetical prote
13	131	5.7	423	2 A33395	acid phosphatase (
14	122.5	5.3	421	2 S14742	acid phosphatase (
15	117	5.1	386	1 JH0610	acid phosphatase (
16	113.5	4.9	344	2 B89130	protein F52E1.8 [i
17	109	4.7	2554	1 TVYF7L	kinase-related pro
18	108.5	4.7	10223	2 T30225	polyketide synthas
19	107.5	4.7	457	2 A56925	paired box transcr
20	107	4.6	450	2 A54429	paired box transcr
21	106.5	4.6	1048	1 BVECS0	exonuclease (EC 3.
22	105.5	4.6	459	2 S22520	paired box transcr
23	104.5	4.5	2535	2 AC0304	probable hemolysin
24	102	4.4	638	2 A29440	signal recognition
25	101.5	4.4	397	2 C81716	hypothetical prote
26	101.5	4.4	479	1 JN0890	acid phosphatase (
27	101.5	4.4	524	2 H75530	probable acid-CoA
28	101.5	4.4	1047	2 C85535	ATP-dependent d5DN
29	101.5	4.4	1047	2 G90684	ATP-dependent d5DN

30	101	4.4	1662	2 T18540	moFA protein precu
31	101	4.4	26926	1 I38344	titin, cardiac mus
32	100.5	4.4	369	2 J02278	hydroxymethylbilan
33	100.5	4.4	438	2 S84682	acid phosphatase (
34	100.5	4.4	478	2 S58315	major plasmid tran
35	100.5	4.4	1214	2 AG2897	conserved hypotet
36	100.5	4.4	1387	2 A97673	probable periplasm
37	100.5	4.4	6260	2 T30228	polyketide synthas
38	100	4.3	328	2 A83572	pyridoxal phosphat
39	100	4.3	1374	1 VCBE17	major capsid prote
40	100	4.3	4135	2 T42629	tenascin-X - bovin
41	99	4.3	636	2 G81983	pinin glycosylato
42	98.5	4.3	425	2 D83186	hypothetical prote
43	98.5	4.3	1034	2 A80551	exonuclease sbcC I
44	97.5	4.2	578	2 H75256	malate oxidoreduct
45	97	4.2	440	2 T32457	hypothetical prote

ALIGNMENTS

RESULT 1

B36733
acid phosphatase (EC 3.1.3.2) precursor - Escherichia coli (strain K-12)
N/Alternate names: phosphoanhydride phosphohydrolase, periplasmic; phytase 6
C/Species: Escherichia coli
C/Date: 19-Apr-1991 #sequence revision 19-Apr-1991 #text change 09-Jul-2004
C/Accession: B36733; S18018; E64839; A26534; S17960; S33278
R/Dassa, J.; Marck, C.; Boquet, P.L.
J. Bacteriol. 172, 5497-5500, 1990
A/Title: The complete nucleotide sequence of the Escherichia coli gene appA reveals sign
A/Reference number: A36733; MUID:90368616; PMID:2168385
A/Accession: B36733
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-432 <DAS>
A/Cross-references: UNIPROT:P07102; UNIPARC:UPI000004750A; GB:M58708; NID:G145283; PIDN
R/Greiner, R.; Jany, K.D.
Biol. Chem. Hoppe-Seyler 372, 664-665, 1991
A/Title: Characterization of a phytase from Escherichia coli.
A/Reference number: S18018
A/Accession: S18018
A/Molecule type: protein
A/Residues: 23-33 <GRE>
A/Cross-references: UNIPARC:UPI0000179815
R/Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A/Title: The complete genome sequence of Escherichia coli K-12.
A/Reference number: A64720; MUID:97426617; PMID:9278503
A/Accession: B64839
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-432 <BLAT>
A/Cross-references: UNIPARC:UPI000004750A; GB:AB0000200; GB:U00096; NID:G2367111; PIDN:AA
A/Experimental source: strain K-12, substrain MG1655
R/Touati, E.; Danchin, A.
Biochimie 69, 215-221, 1987
A/Title: The structure of the promoter and amino terminal region of the pH 2.5 acid phos
A/Reference number: A26534; MUID:87271766; PMID:3038201
A/Accession: A26534
A/Molecule type: DNA
A/Residues: 1-50, 'NACGCHRRWANTAG', 65, 'T', 67-74, 'DV', 77-111, 'S', <TOU>
A/Cross-references: UNIPARC:UPI000016EDE2; GB:X05471; NID:G40925; PIDN:CAA93031.1; PID:G
R/Dassa, J.; Faith, H.; Marck, C.; Dion, M.; Kieffer-Bontemps, M.; Boquet, P.L.
Mol. Gen. Genet. 229, 341-352, 1991
A/Title: A new oxygen-regulated operon in Escherichia coli comprises the genes for a put
A/Reference number: S17958; MUID:92049231; PMID:1658595
A/Accession: S17960
A/Status: not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-17 <DA2>
A/Cross-references: UNIPARC:UPI000016PFB1; GB:S63811; NID:G238656; PIDN:AAB20286.1; PID:

R;Greiner, R.; Konietzky, U.; Jany, K.D.
Arch. Biochem. Biophys. 303, 107-113, 1993
A;Title: Purification and characterization of two phytases from *Escherichia coli*.
A;Reference number: S33278; MUID:93256556; PMID:8387749
A;Accession: S33278
A;Molecule type: protein
A;Residues: 23-31, 'A', 33-34 <GR2>
A;Cross-references: UNIPARC:UPI0000179816
C;Comment: In addition to cAMP-mediated control, this enzyme is induced when bacterial cAMP is present under aerobic conditions.
C;Genetics:
A;Gene: appA
A;Map position: 25 min
C;Keywords: monomer; periplasmic space; phosphohistidine; phosphoprotein; phosphoric monophosphate
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-112/Product: acid phosphatase #status predicted <MAT>
F;38/Active site: Arg #status predicted
F;39/Active site: His (phosphohistidine intermediate) #status predicted

Query Match 97.1%; Score 2235; DB 2; Length 432;
Best Local Similarity 99.5%; Pred. No. 1.2e-171;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKAILIPFLSLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
Db 1 MKAILIPFLSLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDAMP 60

Qy 61 TWPVKLGLWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPSQSGQVAIADVDERTRKTGE 120
Db 61 TWPVKLGLWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPSQSGQVAIADVDERTRKTGE 120

Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAIISRAGGSIAIDFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAIISRAGGSIAIDFTGH 180

Qy 181 RQTAFRELRLVLPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 RQTAFRELRLVLPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy 241 EIFLLOQAQMPGPGWGRITDISHQNTLLSLHNAQFYLLQRTPEVARSRATPLDLIMAA 300
Db 241 EIFLLOQAQMPGPGWGRITDISHQNTLLSLHNAQFYLLQRTPEVARSRATPLDLIMAA 300

Qy 301 LTPHPKQKQAVGLTPTSVLFIAGHDNTLANLGGALELNTLPGQDNTPPGGELVFERW 360
Db 301 LTPHPKQKQAVGLTPTSVLFIAGHDNTLANLGGALELNTLPGQDNTPPGGELVFERW 360

Qy 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTTPGVEVKLTLAGCEERNAQMCSLAGFTQ 420
Db 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTTPGVEVKLTLAGCEERNAQMCSLAGFTQ 420

Qy 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432

RESULT 2
H90770
phosphoanhydride phosphorylase [imported] - *Escherichia coli* (strain O157:H7, substrain H90770)
C;Species: *Escherichia coli*
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: H90770
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic map of enterohemorrhagic *E. coli* O157:H7
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: H90770
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-434 <HAY>
A;Cross-references: UNIPROT:Q8XC29; UNIPARC:UPI000000D0517; GB:BA000007; PIDN:BA034559.1;
A;Experimental source: strain O157:H7, substrain RIMD 050952

C;Genetics:
A;Gene: EC81136

Query Match 96.0%; Score 2210; DB 2; Length 434;
Best Local Similarity 98.4%; Pred. No. 1.2e-169;
Matches 427; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

Qy 1 MKAILIPFLSLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDA 58
Db 1 MKAILIPFLSLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDA 60

Qy 59 WPTWPKLGLWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPSQSGQVAIADVDERTRKT 118
Db 61 WPNWPKLGLWLTTPRGGLIAYLGHYQRLVADGLLTKKGCPOGQVAIADVDERTRKT 120

Qy 119 GEAFAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAISRAGGSIAIDFT 178
Db 121 GEAFAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAISRAGGSIAIDFT 180

Qy 179 GHRQTAFRELRLVLPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASML 238
Db 181 GHRQTAFRELRLVLPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASML 240

Qy 239 LTFEILLOQAQMPGPGWGRITDISHQNTLLSLHNAQFYLLQRTPEVARSRATPLDLIM 298
Db 241 LTFEILLOQAQMPGPGWGRITDISHQNTLLSLHNAQFYLLQRTPEVARSRATPLDLIM 300

Qy 299 AALTTPHPKQKQAVGLTPTSVLFIAGHDNTLANLGGALELNTLPGQDNTPPGGELVFE 358
Db 301 IALTTPHPKQKQAVGLTPTSVLFIAGHDNTLANLGGALELNTLPGQDNTPPGGELVFE 360

Qy 359 RWRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTTPGVEVKLTLAGCEERNAQMCSLAGF 418
Db 361 RWRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTTPGVEVKLTLAGCEERNAQMCSLAGF 420

Qy 419 TQIVNEARIPACSL 432
Db 421 TQIVNEARIPACSL 434

RESULT 3
D85633
hypothetical protein appA [imported] - *Escherichia coli* (strain O157:H7, substrain EDL933)
C;Species: *Escherichia coli*
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: D85633
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, M.W.; Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: D85633
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-444 <STO>
A;Cross-references: UNIPROT:Q8XC29; UNIPARC:UPI0000165751; GB:AE005174; NID:gl2514245; P; J
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: appA

Query Match 96.0%; Score 2210; DB 2; Length 444;
Best Local Similarity 98.4%; Pred. No. 1.2e-169;
Matches 427; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

Qy 1 MKAILIPFLSLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDA 58
Db 11 MKAILIPFLSLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDA 70

Qy 59 WPTWPKLGLWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPSQSGQVAIADVDERTRKT 118
Db 71 WPNWPKLGLWLTTPRGGLIAYLGHYQRLVADGLLTKKGCPOGQVAIADVDERTRKT 130

Qy 119 GEAFAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAISRAGGSIAIDFT 178

Db 131 GEAPAGLAPCAITVHTQADTSSPDPLFNKTVGCQDNLNANVTDAILSRAGGSADFT 190
Qy 179 GHRQTAFRELRLVNFPSNCLCKREKQDECSLTQALPSELKVSADNVSLTGAVSLASM 238
Db 191 GHRQTAFRELRLVNFPSNCLCKREKQDECSLTQALPSELKVSADNVSLTGAVSLASM 250
Qy 239 LTFIFLLQQAQMPGPGGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIM 298
Db 251 LTFIFLLQQAQMPGPGGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIM 310
Qy 299 AALTPHPKQKAYGYTLTSLVLFITAGHDTNLNIGGALSLNWTLPQGDNTPPGGELVPE 358
Db 311 IALTTPHPKQKAYGYTLTSLVLFITAGHDTNLNIGGALSLNWTLPQGDNTPPGGELVPE 370
Qy 359 RWRLSDNSQWIOVSLVFQTLQQRMDKTPLSLNTPPGSEVVKLTLAGCEERNAQGCMSLAGF 418
Db 371 RWRLSDNSQWIOVSLVFQTLQQRMDKTPLSLNTPPGSEVVKLTLAGCEERNAQGCMSLAGF 430
Qy 419 TQIVNEARIPACSL 432
Db 431 TQIVNEARIPACSL 444

RESULT 4
AC0201
acid phosphatase [EC 3.1.3.2] [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AC0201
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AC0201
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-441 <KUR>
A;Cross-references: UNIPROT:Q8ZFP6; UNIPARC:UPI00000DCCF7; GB:AL590842; PIDN:CAC90470.1;
C;Genetics:
A;Gene: YP01648
C;Keywords: phosphoric monoester hydrolase

Query Match 41.0%; Score 943.5; DB 2; Length 441;
Best Local Similarity 45.6%; Pred. No. 6.5e-68;
Matches 199; Conservative 66; Mismatches 156; Indels 15; Gaps 7;

Qy 5 LIPFLSLILPITQSAFAQSEPE-LKLESVVIVSRHGVRAPTKATQLMQDVTTPDAWTPW 63
Db 13 LVLMLSGLAITAPVA--AEPSGYTLERVVILSRHGVRSPTKTQTLMDVTPDKWQWP 69
Qy 64 VKLGWLTFRGELLAYLGHVQRORLVADGLIAKKGCPQSGQVAVIADVDRTRKTGEAFAGLAPDC 123
Db 70 VKAGYLTPRGAELVTLMGGFYGYDFRSLGLLA-AGCPAEGGVYAAQADIDORTLTGQAF 128
Qy 124 AGLAPDCAITVHTQADTSSPDPLFNKTVGCQDNLNANVTDAILSRAGGSADFTGHRQT 183
Db 129 DGVAPEGGLTVHNOADLKKTDPLFHPVEAGVCKLDAQAQTOKAIEQUGGGLDVTVSQRYAK 188
Qy 184 AFRELRLVNFPSNCLCKREKQDECSLTQALPSELKVSAD--NVSLTGAVSLASMLTE 241
Db 189 PFAQMGDVLPFAASPYCKSLQQQKCTCDFAHFAANEVNVNKEGTVLSGFLALSSTLGE 248
Qy 242 IFLLQQAQMPGPGGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 301
Db 249 IFLLQQAQMPGPGGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 308
Qy 302 TPHPKQKAYGYTLTSLVLFITAGHDTNLNIGGALSLNWTLPQGDNTPPGGELVPE 357
Db 309 T---LQIDAQCGKLPISAQNRVFLFGHDTNINIANIAGMLGADWQLPBPQDNTPPGGGLVF 365

Qy 358 ERWRRLSDNSQWIOVSLVFQTLQQRMDKTPLSL-NTPPGEVVKLTLAGCEERNAQGCMSLA 416
Db 366 ELAQNPEDNHQRYAVKMFYQTMQDLRNAEKLDLKNNPAGIISVAVAGCENNGDDKLCELD 425
Qy 417 GFTQIVNEARIPACSL 432
Db 426 TFQKKVAKVIEPACHI 441

RESULT 5
F90773
periplasmic glucose-1-phosphatase [imported] - Escherichia coli (strain O157:H7, substra
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: F90773
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyana, K.; Han, C.G.; Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H. DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F90773
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-413 <HAY>
A;Cross-references: UNIPROT:Q8XBZ6; UNIPARC:UPI00000D0533; GB:BA000007; PIDN:BA034581.1;
A;Experimental source: strain O157:H7, substrain RMD 0509952
C;Genetics:
A;Gene: ECs1158

Query Match 23.7%; Score 545; DB 2; Length 413;
Best Local Similarity 32.1%; Pred. No. 6.1e-36;
Matches 135; Conservative 76; Mismatches 180; Indels 30; Gaps 10;

Qy 13 IPLTPQSAFAQSEPE-LKLESVVIVSRHGVRAPTKATQLMQDVTTPDAWTPWPKLGLMT 70
Db 13 VLLASNAQAQTVPEGYQLOQVLMMSRHLRAPLANNGSVLEQSTPNKNPEWDVPGGOLT 72
Qy 71 PRGGELIAYLGHVQRORLVADGLIAKKGCPQSGQVAVIADVDRTRKTGEAFAGLAPDC 130
Db 73 TKGGVLEVYGHVYRHLAQSGVMSKGECPDPTVYAYANSLQRTVATQAFFITGAPGC 132
Qy 131 AITVHTQADTSSPDPLFNKTVGCQDNLNANVTDAILSRAGGSADFTGHRQTAFREL 190
Db 133 DIPVHQEKMGTMPTFNPVITD---DAAFSEQAAMAEKLSKL--QLTDSVQLLEK 186
Qy 191 VLNPPQSNCLCKREKQDECSLTQALPSELKVSADNVSLTGAVSLASMLTEIFLLQQAQ 250
Db 187 IVNYKDSAPC---KEKQCSLVGDKNTFSKATYQEEFSGVSGPLKVGNSLVDAFTLQYVEG 242
Qy 251 MP--EPGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAAALTPHPQK 308
Db 243 FPMQDVANGIKSDQKVKLSKXNGYQDSLFTSPVARNVAKPLVSYIDKAL----- 295
Qy 309 QAYGVTLPTS---VLFITAGHDTNLNIGGALSLN-WTLPQGDNTPPGGELVFERWRRLS 364
Db 296 ---VTDRTSAPKITVLVGHDSNIALSLTALDFKPYQLHQNERTPIGKIVFQWRD 351
Qy 365 DNSQWIOVSLVFQTLQQRMDKTPLSLNTPPGSEVVKLTLAGCEERNAQGCMSLAGTQIVNE 424
Db 352 ANRLMKIEVYQASQRLRNADALTQAPAQRTVLELSGC-PIDADGFCPMDKFDVSLNE 410
Qy 425 A 425
Db 411 A 411

RESULT 6
B85636
periplasmic glucose-1-phosphatase [imported] - Escherichia coli (strain O157:H7, substra
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: B85636
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

Query Match 23.5%; Score 541; DB 2; Length 417;

Query Match 23.4%; Score 539; DB 2; Length 413;
Best Local Similarity 32.1%; Pred. No. 1.8e-35;
Matches 135; Conservative 75; Mismatches 181; Indels 30; Gaps 10;

QY 13 IPLTQSAFAQSEPE-LKLESVIVSRHGVRAP-TKATQLMQDVTDPDAMPVVKLGMLT 70
DB 13 IYLLASNAQAQVTPPEGYQLQVLMMSRHLRAPLANNGSVLEQSTFNKWPENWDPVGGOLT 72

QY 71 PRGGELIAYLGHYQORLQVADGLLAKKGCQSQGVVAIADVDERTKRTGFAAAGLAPDC 130
DB 73 TKGGLVETMGHYMEWLAEOQWVSGECPPTVTVAYANSQRTVATAQFFITGAFFGC 132

QY 131 AITVHTQADTSSPDLFNLPLKTVGCQLDNANVTDAILSRAGGSIADFTGHRQTAFRELER 190
DB 133 DIPVHHQKMGWMDTFNPFVITD-----DAAAFSEQAVAAMEKLSKL--QLTDSVQLLEK 186

QY 191 VLNFPQSNLCLKREKQDESCSLTQALPSBLKVSADNVSLTGAVSLASMLTEIFLFLQQAQ 250
DB 187 IVNYKDSFAC-----KEKQCSLVGDKNTFSAKYQOQEPGVSGPLKVGNSLVDAFTLQYYEG 242

QY 251 MP--EPGWRITDHSQWNTLLSHNAQFVLLQRTPEVARSRATPLLDLIMAAALTPHPQK 308
DB 243 FPMQVANGETKSDQWVSKLNGYQDSLFTSPEVARNAVAKPLVSDIXAL----- 295

QY 309 QAYGVTLPST---VLFIAGHDTNLNAGALELN-WTLPGQPDNTPPGGELVFERWRRLS 364
DB 296 ---VTDRTSAPKITVLVGHDSNIASLLTALDFKPYQLHDQWERTPIGGKIVFQRWHD 351

QY 365 DNSQWIVSVFQTLQMRDKTPLSLNTPPGEVKLTLAGCEBNAQCMCSLAGFTQIVNE 424
DB 352 ANRDLMKIEYVYQSAEQLRNADALTLOAPQRVTLLELSCG-PIDADGFCPMDKFDVSLNE 410

QY 425 A 425
DB 411 A 411

RESULT 9
EB7316
periplasmic phosphoanhydride phosphohydrolase [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C/Accession: EB7316
R.Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Esmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; MUID:21173698; PMID:11259647
A/Accession: EB7316
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-414 <STO>
A/Cross-references: UNIPROT:Q9AAQ4; UNIPARC:UPI00000C70A0; GB:AE005673; NID:g13421731; E
C/Genetics:
A/Gene: CC0542

Query Match 22.8%; Score 526; DB 2; Length 414;
Best Local Similarity 34.7%; Pred. No. 2.1e-34;
Matches 148; Conservative 57; Mismatches 193; Indels 28; Gaps 12;

QY 9 LSL---LPLTPQSAFAQSEPELKLSEVVIVSRHGVRAPTKATQLMQDVTDPDAMPVVK 65
DB 3 LSLRAVTVTLALASGAASAE--TLEKVVILSRHGVRSAMSPEKLEASARPFRFVP 60

QY 66 LGWLTPRGELIAYLGHYQORLQVADGLLAKKGCQSQGVVAIADVDERTKRTGFAAAG 125
DB 61 AGHLTARGETLVARMGDIYRRHYAAQGLKPCDC---ASVYAWANVTQRTIATAKAYET 117

QY 126 LAPDCAITVHTQADTSSPDLFNLPLKTVGCQLDNANVTDAILSRAGGSIADFTGHRQTAF 185
DB 118 LAPGCFVTNTVGE--GNIDPMFEPYKAGIVKADHALARAAGVGVGGDLTAWASHNOEA 176

QY 186 RELERVNLFPQSNLCL----KREKQDESCSLTQALPSBLKVSADNVSLTGAVSLASMLTE 241
DB 177 EQLDALLMQCKGCPPPAPGRVFDKAPGVFDG--EEL-----AGLSGEAFASGVTE 228

QY 242 IFLLQQAQMGPEPGW-RITDSHOWNTLLSHNAQFVLLQRTPEVARSRATPLDLINAA 300
DB 229 SLLMAWADGRDFAGLGMKSLDEEALTRSFLLHQAEDFLRLRTPYVARTLAGHLADRLAAT 288

QY 301 LTPHPQKQAQGVTLPTSVLFIAGHDTNLNAGALELNWTLPG-OPDNTPPGGELVFER 359
DB 289 LR---DGAAG-IPVDARLVIIAGHDGTLASLGGLLRMEWTLPGYQPNQIQPGGALVFER 344

QY 360 WRRLSDNSQWIVFQTLQMRDKTPLSLNTPPGEVKLTLAGCEBNAQCMCSLAGFT 419
DB 345 WRR--DDGVRVVRFTGQSLSLRNLMTALDAKTPLSAPVFVQSGGTATPAFDCLDEFE 403

QY 420 QIVNEA 425
DB 404 TVVRGA 409

RESULT 10
AG0632
glucose-1-phosphatase precursor (G1Pase), secreted [imported] - Salmonella enterica subsp.
C/Species: Salmonella enterica subsp. enterica serovar Typhi
A/Note: this species has also been called Salmonella typhi
C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C/Accession: AG0632
R.Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, P.; M.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A/Reference number: AB0502; MUID:21534947; PMID:11677608
A/Accession: AG0632
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-413 <PAR>
A/Cross-references: UNIPARC:UPI0000059FE6; GB:AL513382; PIDN:CAD08242.1; PID:g16502289;
C/Genetics:
A/Gene: STV1153

Query Match 22.8%; Score 524; DB 2; Length 413;
Best Local Similarity 33.0%; Pred. No. 3e-34;
Matches 135; Conservative 65; Mismatches 185; Indels 24; Gaps 10;

QY 22 AQSEPE-LKLESVIVSRHGVRAP-TKATQLMQDVTDPDAMPVVKLGWLTPRGELIAY 79
DB 22 AQTTPGVLQVQLVLMMSRHLRAPLANNGVLAQSTPNAPAWPDVPGQLTKGGVLEVY 81

QY 80 LGHYQORLQVADGLLAKKGCQSQGVVAIADVDERTKRTGFAAAGLAPDCAITVHTQAD 139
DB 82 MGHYTREWLVAQGLIPSECEPADPTVTVAYANSQRTVATAQFFITSPAGGCDIPVHGEK 141

QY 140 TSSPDLFNLPLKTVGCQLDNANVTDAILSRAGGSIADFTGHRQTAFRELERVNLFPQSNL 199
DB 142 MGTMDPTFNPFVITDSDAAFRQQAQVA-MEKARSQ-----HLDSYKLEQLITHYQDPS 195

QY 200 CLKREKQDESCSLTQALPSBLKVSADNVSLTGAVSLASMLTEIFLLQQAQCMPE--EPGWG 257
DB 196 C----KEKHQCSLLDAKDTFSANYQOEPGVQGLKVGNSLVDAFTLQYYEGFPMQVANG 251

QY 258 RITDSHOWNTLLSHNAQFVLLQRTPEVARSRATPLDLIMAAALTPHPQKQAQGVTLPT 317
DB 252 GIHTDROKWLKSLKNGYQDSLFTSPTVARNVAAPLVKYIDKVLV-----AERVSAF- 303

QY 318 SVLFIAGHDTNLNAGALELN-WTLPGQPDNTPPGGELVFERWRRLSDNSQWIVSVL 376
DB 304 KVTVLVGHDSNIASLLTALDFKPYQLHDQWERTPIGGQLVFORWHDGNANRDLMKIEYV 363

QY 377 QTLQMRDKTPLSLNTPPGEVKLTLAGCEBNAQCMCSLAGFTQIVNEA 425

Db 364 QSARQLRNAEALTLKSPAQRVTLLEKGPV-DANGFCPLDKFDNMNTA 411

RESULT 11

s06167

acid phosphatase (EC 3.1.3.2) ACP2 precursor [validated] - human

N;Alternate names: acid phosphatase, lysosomal

C;Species: Homo sapiens (man)

C;Date: 04-Dec-1992 #sequence_revision 01-Dec-1995 #text_change 09-Jul-2004

C;Accession: S06167; S05525; S01155

R;von Figura, K.

submitted to the EMBL Data Library, June 1989

A;Reference number: S06167

A;Accession: S06167

A;Molecule type: DNA

A;Residues: 1-423 <VON>

A;Cross-references: UNIPROT:P11117; UNIPARC:UPI0000131FE1; EMBL:X15525; NID:G34239; PIDN:R;Geier, C.; von Figura, K.; Pohlmann, R.

Eur. J. Biochem. 183, 611-616, 1989

A;Title: Structure of the human lysosomal acid phosphatase gene.

A;Reference number: S05525; MUID:89377828; PMID:2776754

A;Accession: S05525

A;Molecule type: DNA

A;Residues: 1-29 <GEI>

A;Cross-references: UNIPARC:UPI00001727CD

R;Pohlmann, R.; Krentler, C.; Schmidt, B.; Schroeder, W.; Lorkowski, G.; Culley, J.; Mer

EMBO J. 7, 2343-2350, 1988

A;Title: Human lysosomal acid phosphatase: cloning, expression and chromosomal assignment

A;Reference number: S01155; MUID:89052645; PMID:3191910

A;Accession: S01155

A;Molecule type: mRNA

A;Residues: 1-423 <POH>

A;Cross-references: UNIPARC:UPI0000131FE1; EMBL:X12548; NID:G34262; PIDN:CAA1064.1; PIDN

A;Note: part of this sequence, including the amino end of the mature protein, was confir

C;Genetics:

A;Gene: GDB:ACP2

A;Cross-references: GDB:118963; OMIM:171650

A;Map position: 11p11.2-11p11.11

A;Introns: 38/3; 70/3; 99/3; 150/3; 183/3; 213/3; 258/1; 285/3; 321/2; 380/1

C;Function:

A;Description: catalyzes the hydrolysis of a wide range of phosphate esters

C;Superfamily: mammalian acid phosphatase

C;Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydrola

F;1-30/Domain: signal sequence #status predicted <SIG>

F;31-423/Product: acid phosphatase ACP2 #status experimental <MAT>

F;41/Active site: Arg #status predicted

F;42/Active site: His (phosphohistidine intermediate) #status predicted

F;92,133,167,177,191,267,322,331/Binding site: carbohydrate (Asn) (covalent) #status pre

F;159-370,212-310,345-349/Disulfide bonds: #status predicted

Query Match 6.1%; Score 139.5; DB 1; Length 423;

Best Local Similarity 23.1%; Pred. No. 0.0023;

Matches 98; Conservative 44; Mismatches 180; Indels 103; Gaps 19;

Qy 30 LESVIVSRHGVRAPTKATQMQDVTDAWPTWPKLWLTTPRGELIAYLGHYQORLV 89

Db 33 LRFVTLIRHGRDSPVKT-----YKDPYQEEWPPQGFQULTKEGMLQHWELGQALRQRY- 87

Qy 90 ADGLLAKGCPQSQQVAIIADVD-ERTRKTGEAPAGLAPDCAITVHTQADTSSPDLP-- 146

Db 88 -HGFL---NTSYHRQEVVVRSTDFRITLMSAEANLAGLFP-----PNGMQR 129

Qy 147 FNP-LKTVGCOLDNANVTDAILSRAGSIADFTGHRQTAFRELVLNFPQSNLCRKREK 205

Db 130 FNPNIWQPIPVHTVPITE-----DRLLKFPPLGPPCPRYEQL 165

Qy 206 QDESCSLTOALPSELKYSADNVSLUTGAVSLASMLTEIFL-----LQQAQGMPEP 254

Db 166 QNE-----TRQTFEYQNESSRNAQFLDMVANETGLDITLETVMNVYDTLFCQTHGLRLP 221

Qy 255 GWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAALTTPHPQKQAYGVT 314

Db 222 PWASPTQWRLSRLLKDFRFLFGIYQQAQKARLQGGVLLAQIRKNTLTMTATTSQ----- 276

Qy 315 LPTSVLPIAGHDTNLNLGGALELNWTLPGQPDNTPPGGELVFERWRLSDNSQWIOVSL 374

Db 277 LP-KLLVYSAHDTTLVALQWALDVY-----NGEQAPYASCHIFELYQEDSGN---FSVEM 327

Qy 375 VFOTLQOQRDKTPLSLNTPPGEVKLTLAGCEENNAQCMCSLAGTQTIVNEARIP-----A 429

Db 328 YF---RNESDKAWPLSLP-----GCPHR-----CFLQDFLRL-TEPVVPKQWQOE 369

Qy 430 CSLRS 434

Db 370 CQLAS 374

RESULT 12

T16058

hypothetical protein F13D11.1 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C;Accession: T16058

R;Fulton, L.

submitted to the EMBL Data Library, November 1995

A;Description: The sequence of C. elegans cosmid F13D11.

A;Reference number: S69020

A;Accession: T16058

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-416 <FUL>

A;Cross-references: UNIPROT:Q19390; UNIPARC:UPI000007AD42; EMBL:U40939; NID:g1072175; P11

C;Genetics:

A;Gene: CESP:F13D11.1

A;Introns: 23/3; 59/3; 88/3; 138/3; 202/3; 244/1; 270/3; 312/3; 349/3; 377/1

C;Superfamily: mammalian acid phosphatase

Query Match 5.8%; Score 133; DB 2; Length 416;

Best Local Similarity 23.4%; Pred. No. 0.0076;

Matches 104; Conservative 60; Mismatches 174; Indels 106; Gaps 24;

Qy 1 MKAILPFFSLIPLTPQSNAPQSEBELKLESVVIVSRHGVRAPTKATQMQDVTDP--- 57

Db 1 MRLVLLLF--FLFPV---AFG-----XLKFVQIWRHGERTPGHLYXFPFGDDLNNVDY 48

Qy 58 ---AWPTWPKLWLTTPRGELIAYLGHYQORLVADGLLAKGCPQSQQVAIIADVDER 114

Db 49 QQIAMP-----CELTGRGILEEFQIG--QLAKIYGEHFGDTYQPRDFHV--YTGKDNR 98

Qy 115 TRKTGEAPAGLAPDCAITVHTQADTSSPDLPFNPLKTGVCOLDNANVTDAILSRAGSI 174

Db 99 TSASAQAMFAGFLPP-----NEDQTNVVELKWQP-----VAQLTDBSIDWVSL---GAI 144

Qy 175 --ADFTGHRQTAFRELVLNFPQSNLCRKREKODESCSLTOALPSELKYSADNVSLTGA 232

Db 145 DNCVPVYGEAQRKSSEYAEVND-----QMEKYD--AELLQLVRNH-----ADE-PIVEA 189

Qy 233 VSLASMLTEI---FLQQAQGMPEPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSR 289

Db 190 VKNHWIDSLKVYIILQDDR-LPYEPWARGENRLNMSFLIHDA-----VVKVQ 238

Qy 290 ATPLLD-----LIMAALTTPHPQKQAYGVTLPSTVLFITAGHDTNLNAGGALELNWTLPGQ 345

Db 239 NDSVGVDYHNLVMSYPETHLQKNSTKGV-----FTSGHDTNLVTIWESLRLD-----GH 287

Qy 346 PDNTPP-GGELVFERWRLSDNSQWIOVSLVFTQLOMRDKTPLSLNTPPGEVKLTLAGC 404

Db 288 PEDIPNYGAHIAIE-----MHBPVQGLSIKFFLSMGFNQTRVELHPHFC 331

Qy 405 EERNAQCMCSLAGTQTIVNEARIP 428

Db 332 -SRSQNDCTWDEFRLLVKSRKP 354

RESULT 13

A33395

acid phosphatase (EC 3.1.3.2) precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 29-Jan-1990 #sequence_revision 29-Jan-1990 #text_change 09-Jul-2004
C;Accession: A33395
R;Himeno, M.; Fujita, H.; Noguichi, Y.; Kono, A.; Kato, K.
Biochem. Biophys. Res. Commun. 162, 1044-1053, 1989
A;Title: Isolation and sequencing of a cDNA clone encoding acid phosphatase in rat liver
A;Reference number: A33395; MUID:89350910; PMID:2764916
A;Accession: A33395
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-423 <HM>
A;Cross-references: UNIPROT:P20611; UNIPARC:UPI0000131FE2; GB:M27893; NID:G202933; PIDN:
C;Superfamily: mammalian acid phosphatase
C;Keywords: phosphohistidine; phosphoprotein; phosphoric monoester hydrolase
F;41/Active site: Arg #status predicted
F;42/Active site: His (phosphohistidine intermediate) #status predicted

Query Match 5.7%; Score 131; DB 2; Length 423;
Best Local Similarity 22.8%; Pred. NO. 0.011;
Matches 105; Conservative 49; Mismatches 184; Indels 122; Gaps 23;

QY 2 KAILIPFL-SLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAW 59
Db 10 QAAQLQFLGLMCLVMPPPIQA-----RSLRFVTLRYHGRDRSPVKA-----YPKDPYQE 58
QY 60 PTWPVKGLMTPRGSELIAVLGHYQORQLVADGLLAKKGCPSQGOVAIIADVD-ERTRKT 118
Db 59 EKNFQGFQKLTKEGMLQHWELGQALRQRY--HGFL---NASHYRQEVYVTRSTDFRDLMS 113
QY 119 GEAFAPAGLAPCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIA 178
Db 114 AEANLAGLFPF-----TEVQHFNISWQPI-----PVHTVPITE----- 148
QY 179 GHRTAFRELRLVNFPOSNCLKREKQDECSSTQALPSELKVSADNVSLTGAVSLASM 238
Db 149 -----DRLKFFLPGPCPRYEQLQNE---TROTPFYQNRSIQNAQFLDMVANETG 194
QY 239 LTEIFL-----LQQAQGMPEPCGWGRITDSHOWNTLLSLHNAQF-VLLQRTPEVA 286
Db 195 LNWLTETIWNVYDTLFCETQTHGLLPPW---ASPQTVALSQLKDFSLFLFGIHEQVQ 251
QY 287 RSR-----ATPLDLIMAAITPHPPQKQAYGVTLPSTVLFIAGHDTNLANLGGALELN 339
Db 252 KARLQGGVLLAQILKNLTLMATTSQPK-----LLVYSAHDTTLVALQWALNVY 300
QY 340 WTLPGQPDNTPPGGELVFERWRRLSDNSQWISQVSLVFTLOQMRDKTPLSLNTPPGEVKL 399
Db 301 -----NGKQAPYASCHIFELYQ--EDNGNF-SVEMYFRNDSK---KAPW-----PL 340
QY 400 TLAGCEERNAGMCSLAGFTQIVNEARIP-----ACSLRS 434
Db 341 TLPQCPHR-----CPLQDFLRL-TEPVIPKDWKECQLAS 374

RESULT 14
S14742
acid phosphatase (EC 3.1.3.2) precursor - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S14742
R;Geier, C.; von Figura, K.; Pohlmann, R.
Biol. Chem. Hoppe-Seyler 372, 301-304, 1991
A;Title: Molecular cloning of the mouse lysosomal acid phosphatase.
A;Reference number: S14742; MUID:91282986; PMID:2059337
A;Accession: S14742
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-421 <GT>
A;Cross-references: UNIPROT:P24638; UNIPARC:UPI0000027103; EMBL:X57199; NID:G52870; PIDN:
C;Superfamily: mammalian acid phosphatase
C;Keywords: lysosome; phosphoric monoester hydrolase

Query Match 5.3%; Score 122.5; DB 2; Length 421;
Best Local Similarity 21.8%; Pred. No. 0.054;
Matches 97; Conservative 49; Mismatches 191; Indels 107; Gaps 19;

QY 2 KAILIPF-----LSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTP 56
Db 8 QAAQLQFLGLMCLVMPPPIQARS-----LRFVTLRYHGRDRSPVKT-----YPKDP 53
QY 57 DAWPTWPVKGLMTPRGSELIAVLGHYQORQLVADGLLAKKGCPSQGOVAIIADVD-BRT 115
Db 54 YQBEKNFQGFQKLTKEGMLQHWELGQALRQRY--HGFL---NTSYHROEVYVTRSTDFRT 108
QY 116 RKTGEAFAPAGLAPCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSTA 175
Db 109 LMSAEANLAGLFPF-----NEVQHFNISWQPI-----PVHTVPITE----- 146
QY 176 DFTGHRTAFRELRLVNFPOSNCLKREKQDECSSTQALPSELKVSADNVSLTGAVSL 235
Db 147 -----DRLKFFLPGPCPRYEQLQNE---TROTPFYQNRSIQNAQFLDMVAN 189
QY 236 ASMLTEIFL-----LQQAQGMPEPCGWGRITDSHOWNTLLSLHNAQFVLLQRTPE 284
Db 190 ETGLTNVLTETIWNVYDTLFCETQTHGLLPPWASPQTVALSQLKDFSLFLFGIHEQVQ 249
QY 285 VARSRAPELDDLIMAAITPHPPQKQAYGVTLPSTVLFIAGHDTNLANLGGALELNWTLPG 344
Db 250 KARLQGGVLLAQILKNLTLMATTSQF-----PKLLVYSAHDTTLVALQWALNVY----- 298
QY 345 QPDNTPPGGELVFERWRRLSDNSQWISQVSLVFTLOQMRDKTPLSLNTPPGEVKLTLAGC 404
Db 299 NGKQAPYASCHIFELYQ--EDNGNF-SVEMYFRNDSK---KAPWPLILP-----GC 343
QY 405 EERNAGMCSLAGFTQIVNEARIP 428
Db 344 PHR-----CPLQDFLRL-TEPVIP 361

RESULT 15
JH0610
acid phosphatase (EC 3.1.3.2) ACP precursor [validated] - human
N;Alternate names: acid phosphatase, prostatic; orthophosphoric monoester phosphohydrolase
C;Species: Homo sapiens (man)
C;Date: 17-Aug-1992 #sequence_revision 01-Dec-1995 #text_change 09-Jul-2004
C;Accession: JH0610; JS0693; A38608; S01331; A32419; S11147; S38863; S41251; S17042; S42
R;Sharief, F.S.; Li, S.S.L.
Biochem. Biophys. Res. Commun. 184, 1468-1476, 1992
A;Title: Structure of human prostatic acid phosphatase gene.
A;Reference number: JH0610; MUID:92272747; PMID:1375464
A;Accession: JH0610
A;Molecule type: DNA
A;Residues: 1-386 <SHA>
A;Cross-references: UNIPROT:P15309; UNIPARC:UPI0000039832; GB:M97580; GB:M97581; GB:M975
A;Accession: JS0693
A;Molecule type: mRNA
A;Residues: 1-386 <SH3>
A;Cross-references: UNIPARC:UPI0000039832; GB:M97589; NID:q189611; PIDN:AAA60021.1; PID:
R;Van Etten, R.L.; Davidson, R.; Stevis, P.E.; MacArthur, H.; Moore, D.L.
J. Biol. Chem. 266, 2313-2319, 1991
A;Title: Covalent structure, disulfide bonding, and identification of reactive surface a
A;Reference number: A38608; MUID:91115848; PMID:1989985
A;Accession: A38608
A;Molecule type: mRNA
A;Residues: 1-386 <VAN>
A;Cross-references: UNIPARC:UPI0000039832; GB:M34840; NID:q189620; PIDN:AAA69694.1; PID:
R;Viikko, P.; Virkkunen, P.; Henttu, P.; Roiko, K.; Solin, T.; Huhtala, M.L.
FEBS Lett. 236, 275-281, 1988
A;Title: Molecular cloning and sequence analysis of cDNA encoding human prostatic acid p
A;Reference number: S01331; MUID:86312981; PMID:2842184
A;Accession: S01331
A;Molecule type: mRNA
A;Residues: 1-14,'A',16,'ASC',20,'CF',23,'C',25-65,'WIWPTHPA',74-213,'A',213-386 <VIH>
A;Cross-references: UNIPARC:UPI000016A48C; EMBL:X52174; NID:g28321; PIDN:CAA36422.1; PID

A;Note: part of this sequence, including the amino end of the mature protein, was confirmed by Sharief, F.S.; Lee, H.; Leuderman, M.M.; Lundwall, A.; Deaven, L.L.; Lee, C.; Li, S.S. Biochem. Biophys. Res. Commun. 160, 79-86, 1989

A;Title: Human prostatic acid phosphatase: cDNA cloning, gene mapping and protein sequencing

A;Reference number: A32419; MUID:89228054; PMID:2712834

A;Accession: A32419

A;Molecule type: mRNA

A;Residues: 1-14, 'A', 16, 'ASC', 20, 'CF', 23, 'C', 25-94, 'D', 96-115, 'R', 117-214, 'S', 216-293, 'T'

A;Cross-references: UNIPARC:UPI000016AE62; GB:M24902; NID:g189618; PIDN:AA60022.1; PID:R;Note: the authors translated the codons GAC for residue 95 as Glu, Cgt for residue 116

R;Tailor, P.G.; Govindan, M.V.; Patel, P.C. Nucleic Acids Res. 18, 4928, 1990

A;Title: Nucleotide sequence of human prostatic acid phosphatase determined from a full-length cDNA

A;Reference number: S11147; MUID:90370491; PMID:2395659

A;Accession: S11147

A;Status: preliminary; translation not shown

A;Molecule type: mRNA

A;Residues: 1-14, 'AFASC', 20, 'CF', 23, 'C', 25-45, 'H', 47-65, 'RIWPTHPA', 74-138, 'E', 140-156, 'R'

A;Cross-references: UNIPARC:UPI000016AF1A; EMBL:X53605; NID:g35683; PIDN:CAA37673.1; PID:R;Banas, B.; Blaschke, D.; Fittler, F.; Hoerz, W. submitted to the EMBL Data Library, April 1993

A;Description: Characterization of the promoter of the human prostatic acid phosphatase

A;Reference number: S38863

A;Accession: S38863

A;Molecule type: DNA

A;Residues: 1-40 <BA>

A;Cross-references: UNIPARC:UPI00000006BB; EMBL:X711391; NID:g416530; PIDN:CAA50514.1; PID:R;Virkkunen, P.H.; Hedberg, P.; Palvimo, J.J.; Birr, E.; Porvari, K.; Taavittainen, P.; submitted to the EMBL Data Library, September 1993

A;Description: Structural organization of human and rat prostate-specific acid phosphatase sequence in the human gene promoter.

A;Reference number: S41251

A;Accession: S41251

A;Molecule type: DNA

A;Residues: 1-40 <VIR>

A;Cross-references: UNIPARC:UPI00000006BB; EMBL:X74961; NID:g439657; PIDN:CAA52913.1; PID:R;Lee, H.; Chu, T.M.; Li, S.S.L.; Lee, C. Biochem. J. 277, 759-765, 1991

A;Title: Homodimer and heterodimer subunits of human prostate acid phosphatase.

A;Reference number: S17042; MUID:91336999; PMID:1908222

A;Accession: S17042

A;Status: preliminary

A;Molecule type: protein

A;Residues: 33-49 <LEE>

A;Cross-references: UNIPARC:UPI00001727CC

R;Banas, B.; Blaschke, D.; Fittler, F.; Hoerz, W. Biochim. Biophys. Acta 1217, 188-194, 1994

A;Title: Analysis of the promoter of the human prostatic acid phosphatase gene.

A;Reference number: S42730; MUID:94153995; PMID:8110833

A;Accession: S42730

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-40 <BA2>

A;Cross-references: UNIPARC:UPI00000006BB; GB:X71391; NID:g416530; PIDN:CAA50514.1; PID:R;Sharief, F.S.; Li, S.S. Biochem. Mol. Biol. Int. 33, 561-565, 1994

A;Title: Nucleotide sequence of human prostatic acid phosphatase ACP gene.

A;Reference number: I37175; MUID:95038536; PMID:7951074

A;Accession: I37175

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-386 <RES>

A;Cross-references: UNIPARC:UPI0000039832; EMBL:U07097; NID:g515995; PIDN:AA860640.1; PID:R;Comment: This protein is synthesized under androgen regulation by epithelial cells of C;Genetics:

A;Gene: GDB:ACPP

A;Cross-references: GDB:119644; OMIM:171790

A;Map position: 3q21.3-q25.2

A;Introns: 40/3; 72/3; 101/3; 152/3; 185/3; 216/3; 260/3; 288/3; 323/2

C;Function:

A;Description: catalyzes the hydrolysis of a wide range of phosphate esters

C;Superfamily: mammalian acid phosphatase

C;Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydrolase

F;1-32/Domain: signal sequence #status predicted <SIG>

F;33-36/Product: acid phosphatase ACP #status experimental <MAT>

F;43,86/Active site: Arg #status predicted

F;44/Active site: His (phosphohistidine intermediate) #status predicted

F;94,220,333/Binding site: carbonylate (Asn) (covalent) #status experimental

F;161-372,215-313,347-351/Diulfide bonds: #status experimental

Query Match 5.1%; Score 117; DB 1; Length 386;

Best Local Similarity 19.8%; Pred. No. 0.13;

Matches 87; Conservative 63; Mismatches 178; Indels 112; Gaps 15;

QY 3 AILPFLSLIPLTPQSAFAQSEFELKLSBWSVSRHGVRAPTKATQLMODVTPDAMPT- 61

DB 13 SLSLGLFLFLFWLDRSLAK-----ELKFTLVFRHGDRSPI-----DTFPTD 56

QY 62 -----WPVKLGMITPFGGELIAYLGHYQRLVADGLLAKKCPQS---GQVAIIADVD 112

DB 57 PIKESWPQFGQLTQLGMEHQHYELGEYIRKRY-----RKFLNBSYKHEQVYIRSTDV 109

QY 113 ERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLENPLKTVGCOLDNANVTDAISRAGG 172

DB 110 DRTMSAMTNLAALFPPEGVSIV-----NPILLMQPIPVHTVPLSE----- 150

QY 173 SIADFTGHRQTAFRELERVLNFPQSNLCKREKQDECSLTQALPSELKVSADNVSITGA 232

DB 151 -----DQLLYLPPRN-CPRFQLESETLKSBEFQKELHPYKDFIATLTK 193

QY 233 VS-----LASMLTEIFLLOQAQMPGPGHGRITDSHOWNTLLSHNAQFYLLQRT 283

DB 194 LSLGLHQDLFGIWSKVYDPLYCESVHFTLPSWATEDTMTKRELSELSLLSLYGIHKOK 253

QY 284 EVARSRTATPLDLIIMAAALTPHPQKQAYGVTLPSTVLFIAGHDTNLANTLGGALEL-NWTL 342

DB 254 EKSLQGGVLDVNEILNHM-----KKATQIPSYKKLIMYSAHDTTVSGLOMALDVYNGLL 307

QY 343 PGQPDNTTPGGEIVFRRWRRLSDNSQWISLVFQTLQMRDKTLPISLNTTPPGEVKLTLA 402

DB 308 P--PYASCHLTLYFEKGEYF-----VEMYRN-----ETQHEPYPLMLP 345

QY 403 GCERNAQMGCSLAGTQIV 422

DB 346 GCSP-----SCPLERFAELV 360

Search completed: June 13, 2006, 10:29:29

Job time : 28.7339 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2006, 10:46:00 ; Search time 10.0917 Seconds
(without alignments)
551.244 Million cell updates/sec

Title: US-10-601-319-2

Perfect score: 2302

Sequence: 1 MKAILPFLSLILPLTPQSA.....IVNEARIPACSLRSHHHHH 440

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 64916 seqs, 12643201 residues

Total number of hits satisfying chosen parameters: 64916

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:*

- 1: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US09_NEW_PUB.pap:*
- 2: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US06_NEW_PUB.pap:*
- 3: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US07_NEW_PUB.pap:*
- 4: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US08_NEW_PUB.pap:*
- 5: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/PCT_NEW_PUB.pap:*
- 6: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US10_NEW_PUB.pap:*
- 7: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US11_NEW_PUB.pap:*
- 8: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US60_NEW_PUB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98	4.3	1730	7	US-11-270-040-6
2	97	4.2	460	6	US-10-953-349-34904
3	97	4.2	464	6	US-10-953-349-34903
4	97	4.2	472	6	US-10-953-349-34902
5	93.5	4.1	4391	7	US-11-183-325-56
6	91.5	4.0	635	6	US-10-511-937-2424
7	91	4.0	458	6	US-10-953-349-8297
8	91	4.0	613	6	US-10-953-349-8296
9	91	4.0	614	6	US-10-953-349-8295
10	90	3.9	384	7	US-11-293-697-2758
11	87.5	3.8	952	7	US-11-293-697-3284
12	86	3.7	1336	6	US-10-511-937-2564
13	84	3.6	370	6	US-10-196-749-350
14	84	3.6	725	7	US-11-293-697-3258
15	82.5	3.6	360	6	US-10-953-349-22123
16	81.5	3.5	838	7	US-11-295-336-2
17	81	3.5	281	6	US-10-953-349-23762
18	80.5	3.5	5738	6	US-10-505-528-150
19	79.5	3.5	269	6	US-10-953-349-7
20	79.5	3.5	282	6	US-10-953-349-6
21	79	3.4	505	7	US-11-293-697-3134
22	79	3.4	826	7	US-11-301-554-330
23	79	3.4	1193	6	US-10-505-928-537
24	78.5	3.4	765	7	US-11-317-329-3
25	78.5	3.4	765	7	US-11-317-329-34

ALIGNMENTS

RESULT 1

US-11-270-040-6
; Sequence 6, Application US/11270040
; Publication No. US20060110762A1
; GENERAL INFORMATION:
; APPLICANT: Kapi, Sanjay
; APPLICANT: Kim, Jeong-Ki
; TITLE OF INVENTION: PORCINE REPRODUCTIVE AND RESPIRATORY
; SYNDROME VIRUS RECEPTOR COMPONENTS AND USES THEREOF
; FILE REFERENCE: 14337/004001
; CURRENT APPLICATION NUMBER: US/11/270,040
; PRIOR FILING DATE: 2005-11-09
; PRIOR APPLICATION NUMBER: 60/680,297
; PRIOR FILING DATE: 2005-05-12
; PRIOR APPLICATION NUMBER: 60/626,788
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1730
; TYPE: PRT
; ORGANISM: Sus scrofa
US-11-270-040-6

Query Match 4.3%; Score 98; DB 7; Length 1730;
Best Local Similarity 21.2%; Pred. No. 2.1;
Matches 94; Conservative 54; Mismatches 184; Indels 112; Gaps 20;
Qy 22 AQSEPELKESVIVSRHGVRAPTKATQMLMODVTPDAMTPVVKLWLTGRLIAYLG 81
Db 1250 ANTSLELRLEGVQVALAPSATVPEGAPVTTCEDPAARP--PTLYVW----- 1294
Qy 82 HYQORLVADGLAKKCPQSGQVIAIADYDETRKTEGAFAAGLAPDCAITV-HTQAD- 139
Db 1295 -YNSRWLQEGSAASLSFP----AATRAHAGAYTCVQDQAGTRISQNALHLYAPRDA 1349
Qy 140 --TSSPDPLFNPKTGVCLDNANVTDAILSRAGGSADF-----TGHRQTAPREL 188
Db 1350 VLSFWDSPASPMVAVQCTVDSPPAEMTLTSHGKVLATSHGVHGLAVGTGHVQVARNAL 1409
Qy 189 E-RVLNFPQSN-----LCLKREKQDESCSLTQALPSELKVSAD---NVSLTGAVSLASMLT 240
Db 1410 QLRVQNVPSRDKDTTYVCMDRNSLGSVSTMQLQPEGVHVVAEPGLDVEGTALNLSCLRP 1469
Qy 241 EIFLLQQAQGMPEPG-----W---GRITDSHOWNTLLSLHNAQ-----FYLLQRTPE 284
Db 1470 S-----GPGHIGNSTFAWFRNGRLQHTSVPTLTFTTHVARAQAGLYHCQABLP 1518

Db 171 KPDLTVDPDVPPTTVAHRCFEAREL--FKPGLVPDDSGVSEHREFGVSGESQIVGI 228
 Qy 360 WRLSDNSQWISVQVSLVFTLOQMRDK--TPLSLNTPPGGEVXLTLAGCE-----ERNAQG 411
 Db 229 RSRTELESEMLOV-----LGKLYKVPVPIGLRPPPP--PTQDIAGHEATLRLWLDROAQG 280
 Qy 412 MCSLAGPTQIVNEARIPACSLRS 434
 Db 281 SVVYAAFG---SEAKLTSAQLQT 300

RESULT 5
 US-11-183-325-56
 ; Sequence 56, Application US/11183325
 ; Publication No. US20060104898A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Vanderbilt University
 ; APPLICANT: Hallahan, Dennis E
 ; APPLICANT: Qu, Shimin
 ; TITLE OF INVENTION: IN VIVO PANNING FOR LIGANDS TO RADIATION-INDUCED MOLECULES
 ; FILE REFERENCE: 1242/47/2/2 CIP
 ; CURRENT APPLICATION NUMBER: US/11/183,325
 ; CURRENT FILING DATE: 2005-07-15
 ; PRIOR APPLICATION NUMBER: US 60/328123
 ; PRIOR FILING DATE: 2001-10-03
 ; PRIOR APPLICATION NUMBER: US 10/259,087
 ; PRIOR FILING DATE: 2002-09-27
 ; NUMBER OF SEQ ID NOS: 56
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 56
 ; LENGTH: 4391
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; PUBLICATION INFORMATION:
 ; DATABASE ACCESSION NUMBER: P98160
 ; DATABASE ENTRY DATE: 2003-02-28
 ; RELEVANT RESIDUES: (1)..(4391)
 US-11-183-325-56

Query Match 4.1%; Score 93.5; DB 7; Length 4391;
 Best Local Similarity 21.4%; Pred. No. 21;
 Matches 110; Conservative 50; Mismatches 186; Indels 167; Gaps 25;

Qy 6 IPFSLSLIPITPQSAFAQSEBELKLSV-VIVSRHGVRAPTKATQLMQDVTDPDAMPWPV 64
 Db 3389 LPATSIPAGSTPT---VQVTPQLETKSIGASVEFHCAPVSDRGTLQ-----RWFK 3435

Qy 65 KLGLWTPRGELIAYLGHYQORLVADGLLAKKCPQSGQVAIIADVDRTRKTGBAPAA 124
 Db 3436 EGGQLPP-----GH-----SVQDGLVLRIONLDQSCQGYTC---QAHGPMGKAQAS 3478

Qy 125 G-----LAPDCAITVHTQADT-----SSPDP-----LFNPLKTGVCOLDN 159
 Db 3479 AOLVIALPSVLINIRTSVQTVVGVHAFVEFCLAGDKPKQVTVKSGGHLRPIQVQ--- 3535

Qy 160 ANVTDAILSRAGGSIADTGHROTAFRELERVLNFPQSNLCKREKQDECSLTQAL--- 216
 Db 3536 ---SGGVRIAHVELADAGVRCFA---TWAAGTTQSHVLL-----LVQALPQI 3578

Qy 217 ---PSELKVSADNVELTAVSLASMLTEIFLQOQGMPEP---GNGRI-----TDSHQWNT 267
 Db 3579 SMPQEVVRPAGSAAVFPCI-----ASGYPTDISKSLDGSLLPDDSRLENN 3624

Qy 268 LLSL-----HNAQPYLLQRTPEVAR-----SRATPLLDLIMAALTTPHPQKQAY- 311
 Db 3625 MLMLPSVRPDAGTYVCTATNRQKVKAFAPHLQVPRVVFYTTQPTPSFLPLPTIKDAYR 3684

Qy 312 -----GVTLPSTSVLFIAGHDITNLAMLGALLENLWTLPGQPDNTPPG----- 352
 Db 3685 KFEIKITFRPDSADGMLLYNGKQKVPGPSPTNLN-----RQDFIFSLGLVGR 3732

Qy 353 GELVFERKRLSDNSQWISVQL-VFQTLQMRDKTPLSL-----NTPPGEV----- 397

Db 3733 PEPFRDAGSGMATIRHPTPLALGHFHTVTLRLSLTQSSLI VGDLPVNGTSQKFGQGLDL 3792
 Qy 398 --KLTACBEBRNAGMCSL-AGFTQIVNEARI 427
 Db 3793 NEELYLGYPDYGAIPKAGLSSGFIGCVREURI 3825

RESULT 6
 US-10-511-937-2424
 ; Sequence 2424, Application US/10511937
 ; Publication No. US20060088836A1
 ; GENERAL INFORMATION:
 ; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
 ; APPLICANT: Wohlgenuth, Jay
 ; APPLICANT: Fry, Kirk
 ; APPLICANT: Woodward, Robert
 ; APPLICANT: Ly, Ngoc
 ; APPLICANT: Prentice, James
 ; APPLICANT: Morris, MacDonald
 ; APPLICANT: Rosenberg, Steven
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
 ; FILE REFERENCE: 506612000104
 ; CURRENT APPLICATION NUMBER: US/10/511,937
 ; CURRENT FILING DATE: 2004-10-19
 ; PRIOR APPLICATION NUMBER: PCT/US2003/012946
 ; PRIOR FILING DATE: 2003-04-24
 ; PRIOR APPLICATION NUMBER: US 10/131,831
 ; PRIOR FILING DATE: 2002-04-24
 ; PRIOR APPLICATION NUMBER: US 10/325,899
 ; PRIOR FILING DATE: 2002-12-20
 ; NUMBER OF SEQ ID NOS: 3117
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 2424
 ; LENGTH: 635
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-511-937-2424

Query Match 4.0%; Score 91.5; DB 6; Length 635;
 Best Local Similarity 19.8%; Pred. No. 1.9;
 Matches 111; Conservative 65; Mismatches 196; Indels 189; Gaps 25;

Qy 2 KAILIPLSLILPITPQSAFAQSEP-ELKL---ESVVIVS-----RHGVRAPTKAT-- 48
 Db 123 RVLFDVDSVGLPAPPSIITKMGSGQPGELQISWEPAPETISDFLYELRYGRDPDNPSTGP 182

Qy 49 ---QLMQ-----DVTPDAMPWPVKLG-----WLTTPRGGE-L 76
 Db 183 TVIQLIATETCCPALQRPHSALDQSPCAQPTMPWQDPKQTSPSREASALTAEKGSCL 242

Qy 77 IAYL----GHYQORLVADGLLAKKCPQSGQVAIIADVDRTRKTGEAFAGLAPDCAI 132
 Db 243 ISGLQPNQSNYLQIRSEPDG-LSLGGSGWSLSPVTVDL-----PGDAVALGL----- 289

Qy 133 TVHTQADTSSPDPLENPLKTKGVQCLDNANVTDAILSRAGGSIADFTGHROTAFRELERVL 192
 Db 290 ----QCFT-----LDLKNTVCWQQQDHASSQGGFFHRSRARCPRDR-- 327

Qy 193 NFPQSNLCKREKQDECSLTQALPSLKVKSADNV-----SLTGAVSLASMLTEIFLLQ 246
 Db 328 -YPIWENCEBEETNPGLOTPQPSRCHFKSRNDSIIHILVEVTTAPGTVHSLGSPFWIH 386

Qy 247 QAQGMPEPG--WGRITDSH-----QWNTLLSLHNAQF 276
 Db 387 QAVRLPTPNLHWREISISGHLEWHQHPSSWAAQETCYQLRYTGEHQDKWVLEPPLGARG 446

Qy 277 YLLQRTPEVARSRATPLLDLIMAALTTPHP------QKQAVGVTLPSTSVLFIAG 324
 Db 447 GTLELPP---RSRYRLQLRARLNGPTYQGPWSSNSDPTRVETATETATWISLVTALHVLVG 503

Qy 325 HDTNLANLGALLENLWTLPGQ-----PD-----NTPPGELVPE 358

Db 504 ----LSAVLGLLLLRQFPAHYRRLRHAWPSLPDLHRLVGLQYLRDTAALSPPKA----- 554
Qy 359 RWRRLSDNSQWIQVSLVFOTLQMRDKTPLSLNTPPGEV-----KLTAGCEER 407
Db 555 ---TVSDTCBEVPSLL-BILPKSSERTPLPLCSSQAQMDYRRRLQPSCLGTMPLSVCPPM 610
Qy 408 NAQCMCSLAGFTQIVNEARIP 428
Db 611 AEGSCCT---THIANHSYLP 628

RESULT 7
US-10-953-349-8297
; Sequence 8297, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 8297
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-8297

Query Match 4.0%; Score 91; DB 6; Length 458;
Best Local Similarity 20.9%; Pred. No. 1.3;
Matches 50; Conservative 31; Mismatches 92; Indels 66; Gaps 10;

Qy 140 TSSPDPLFNPLKGTGVCOLDNANVTDAILSRAGGSIADF-----TGHROTAFRELERVL 192
Db 216 SGSPNGLLSDIK-----FKNMNLTFRWSNSYAGLDYRPGCQGLVNHNRATSGIIMEHVN 270
Qy 193 NFPQSNCLKREKQDESCSLTQALPSELKVS-ADNVSLTGAVSLASMLTIFILLQQAQGM 251
Db 271 GFRVENVDLK-WSDDDDVNAAMNVPLEFRPSTVNNVSFVGFTS--GLYTKLFESDYVMSL 327
Qy 252 PEPGWRITDISHQWNTLLSLHNAQFYLLQRTPEVARSEATPLDLIMAALTTPHPQKQAY 311
Db 328 -----FTSNHQYIT-----ESIGITKMASTKPLIVAA----- 355
Qy 312 GVTLPSTVLFIAGHDTNLANLGALBNWTLPGQPDN-----TPPGGELVFERW 360
Db 356 -KTLRNRIHFRSGSTSGPS-----RWATPGHEERPKGYFMNRTPPPPGQSRKQWEDW 406

RESULT 8
US-10-953-349-8296
; Sequence 8296, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 8296
; LENGTH: 613
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-8296

Query Match 4.0%; Score 91; DB 6; Length 613;
Best Local Similarity 20.9%; Pred. No. 2;
Matches 50; Conservative 31; Mismatches 92; Indels 66; Gaps 10;

Qy 140 TSSPDPLFNPLKGTGVCOLDNANVTDAILSRAGGSIADF-----TGHROTAFRELERVL 192
Db 371 SGSPNGLLSDIK-----FKNMNLTFRWSNSYAGLDYRPGCQGLVNHNRATSGIIMEHVN 425
Qy 193 NFPQSNCLKREKQDESCSLTQALPSELKVS-ADNVSLTGAVSLASMLTIFILLQQAQGM 251
Db 426 GFRVENVDLK-WSDDDDVNAAMNVPLEFRPSTVNNVSFVGFTS--GLYTKLFESDYVMSL 482
Qy 252 PEPGWRITDISHQWNTLLSLHNAQFYLLQRTPEVARSEATPLDLIMAALTTPHPQKQAY 311
Db 483 -----FTSNHQYIT-----ESIGITKMASTKPLIVAA----- 510
Qy 312 GVTLPSTVLFIAGHDTNLANLGALBNWTLPGQPDN-----TPPGGELVFERW 360
Db 511 -KTLRNRIHFRSGSTSGPS-----RWATPGHEERPKGYFMNRTPPPPGQSRKQWEDW 561

RESULT 9
US-10-953-349-8295
; Sequence 8295, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 8295
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-8295

Query Match 4.0%; Score 91; DB 6; Length 614;
Best Local Similarity 20.9%; Pred. No. 2;
Matches 50; Conservative 31; Mismatches 92; Indels 66; Gaps 10;

Qy 140 TSSPDPLFNPLKGTGVCOLDNANVTDAILSRAGGSIADF-----TGHROTAFRELERVL 192
Db 372 SGSPNGLLSDIK-----FKNMNLTFRWSNSYAGLDYRPGCQGLVNHNRATSGIIMEHVN 426
Qy 193 NFPQSNCLKREKQDESCSLTQALPSELKVS-ADNVSLTGAVSLASMLTIFILLQQAQGM 251
Db 427 GFRVENVDLK-WSDDDDVNAAMNVPLEFRPSTVNNVSFVGFTS--GLYTKLFESDYVMSL 483
Qy 252 PEPGWRITDISHQWNTLLSLHNAQFYLLQRTPEVARSEATPLDLIMAALTTPHPQKQAY 311
Db 484 -----FTSNHQYIT-----ESIGITKMASTKPLIVAA----- 511
Qy 312 GVTLPSTVLFIAGHDTNLANLGALBNWTLPGQPDN-----TPPGGELVFERW 360
Db 512 -KTLRNRIHFRSGSTSGPS-----RWATPGHEERPKGYFMNRTPPPPGQSRKQWEDW 562

RESULT 10
US-11-293-697-2758
; Sequence 2758, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2758

Query Match 3.9%; Score 90; DB 7; Length 384;
Best Local Similarity 24.2%; Pred. No. 1.2;
Matches 66; Conservative 24; Mismatches 95; Indels 88; Gaps 15;

QY 5 LIPPLSLLI-----PLTPOSAPQSPPELKLSESVIVSRHGVRAPTKATQLMQDVTTPDWP 60
DB LLELPQTQISSRPTPRCKSSSRPHL-----ANGRRSPLKAAQ--ODRGLPAWG 117

QY 61 TWPVKLWLTTPRG--GELIAYLGHY-QRQRLVADGLLARK-----GCPQSQQVAII 108
DB VMLLSRAEVRPRASALGSSRRGLDGKPRSRAAEFTQARPTKRTPCGGRPEAG--CWL 175

QY 109 ADVDERTKTGEAFAAGLAPDCAITVHTQADTSSDPLFNPLKTVGCQLDNANVTDAIIS 168
DB ENRRQRRKAG-AERAGAGP-----TLTPPLPAGSPDPF-SPAATAAPQPEQR----PHLS 225

QY 169 RAGGSIADFTGHRQTAFRELERVLNFPQSNLCLKREKODE-----SCSLTQALPSELKVS 223
DB 226 RG-----QWLLPRRPERAAGSAGSCSSRRALPA-----S 256

QY 224 ADNVSITGAVSLASMLTIFLLQQAQGMPEPGW 256
DB 257 GSGAAMTGSWA-----APPRPAW 274

RESULT 11
US-11-293-697-3284
; Sequence 3284, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; PRIOR FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3284
; LENGTH: 952
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-3284

Query Match 3.8%; Score 87.5; DB 7; Length 952;
Best Local Similarity 21.7%; Pred. No. 7.8;
Matches 84; Conservative 50; Mismatches 144; Indels 109; Gaps 19;

QY 8 FLSLL--IPLTPQSAFAQSEPE-LKLESVIVSRHG-----VRAPTQATQLMQDVTTPDA 58
DB 123 FSSLLSDVSIILPNAEEKAGGESVQTSVSAKEHRRRMGKLLRRASK-QWIRDCNPEV 181

QY 59 WPTWPKLWLTTP-----RGELIA-YLGHYQ-----RQRLVA-----DGLL 94
DB 182 SESFKVQDEVFVPTKPDWCDAELAAITSINHKONHLSLVEHQSLQACDDLLPYPHGLL 241

QY 95 AKK-----GCPQSQQVAIIADVDERTK-----TGSAFAAGLAPDCAI 132
DB 242 SKQWPCSSMPAKQAPPSCVSEG-----SVKGRQKENLFTQNKLSRLRAGKGPD--- 292

QY 133 TVHTQADTSDDPLFNPLKTVGCQLDNANVTDAIILSRAGGSIAFTGHRQTAFRELE--RV 191
DB 293 SFEMBEVQSPFEETGNPLDM-----TSGTVCARVDRANSQSDSSGPLEBPLPLQPM 346

QY 192 LNFQSNLCLK--REKQDECSLTQALPSELK-----VSADNVSLTGAVSL 235
DB 347 PSLPNSQSPAENGGRKPRDQSHLSVSSQDCQLESDDGPDSDKSRASMSFSQEVNALEQRAS 406

QY 236 ASMLTEIFLQQAQGMPE-----PGWGRITDTSHOWNTLLSLHNAQFYLLQRT 282
DB 407 VSMEEERFLLEAMEGPPELYIPDMACAKTTRTGECPRKDSHLWQ-LLPMPHAEY----- 459

QY 283 PEVARSRATPLLDLIMAALTTPHPQPKQ 309
DB 460 -EVRPTATSKYDHPPLGFMVTHVTEMQ 485

RESULT 12
US-10-511-937-2564
; Sequence 2564, Application US/10511937
; Publication No. US2006008836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Morris, James
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2564
; LENGTH: 1336
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2564

Query Match 3.7%; Score 86; DB 6; Length 1336;
Best Local Similarity 21.4%; Pred. No. 17;
Matches 99; Conservative 56; Mismatches 163; Indels 144; Gaps 26;

QY 12 LIPLTPOSAPQSAPEPELKLSESVIVSRHGVRA-----TKATQLMQDVTTPDAMPWPKL 66
DB 773 LASLTPTGQTALRGP-LYTSLAQLYSHHGCHGPAITFTMTQAVEASAIAGVRAIVDHLVAL 831

QY 67 GMLTPRGSELIAVLGHYQ--RQRLVA-----DGLLAKKGCPSQGOVAIIADVDERTKTGE 120
DB 832 AMLHVLHGQSPVALDIILQSVRDVAVASEDQEGVIA-----NMVAVALKKTGTRQAAE 884

QY 121 AF-----AAGLAPDCAITVHTQADTSSPD-----PLFNPLKTVGCQOLD 158
DB 885 SYTEALRVARDLQQRNQAVGLANFGALCLHAGASRLAQHVLLEAVRLFSLPLGECGRD 944

QY 159 NANY-----TDAILSRAGS-----IADTGHRTQTAFLRELVNLF-----POSN 198
DB 945 FTHVLLQLHGLCTRQGPAAQGGKGYEWALLVAVEMGHVESQLRAVORLCHFYSAVMPSEA 1004

QY 199 LC-LKREKQ--DECSLT-QALPSELKVSADNVSLT-----CAVSLASMLTIFLL-LQ 246
DB 1005 QCVIYHELQSLSPACKVADKVLLEGQLLETISQLYLSLTERAYKSAIDYTKRSLGIFIDLQ 1064

QY 247 QAQGMPEFGWGRITDTSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAALTTPHP 306
DB 1065 KKE-----KEAHAW-----LQAGKIYYILLRQSE-----LVDLYI----- 1093

QY 307 QKQAYGVTLSVFLFIAGHDTNLANLGGALELWNTLPGQDNTPPGGELVFE--RWRLUSD 365
DB 1094 -QVAQNVALYTG-----DPNL-----GLEL-----FEAAGDIFDGAWER--- 1127

Qy 366 NSQWQVSLVFTLQQMRDXT-PLSLNTPPGEVKLTLAGCEE 406
Db 1128 -----EKAVSFYDRALPLAVTT--GNRKAEILCNK 1157

RESULT 13

US-10-196-749-350
; Sequence 350, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; CURRENT FILING DATE: 2002-07-16
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 350
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-196-749-350

Query Match 3.6%; Score 84; DB 6; Length 370;
Best Local Similarity 22.9%; Pred. No. 4.1;
Matches 61; Conservative 38; Mismatches 123; Indels 44; Gaps 12;
Qy 68 WLTPRGGLIAYLGHYQRLVADGLLAKGCPQSGOVAIADVDERTKRTGEAFAAGLA 127
Db 86 WLKP---AMTQAQDEVQGERLSEARLSQRLDSTAEADLSDFEE-CEETGELFEFP-A 140
Qy 128 PDCAITVHTQADTSSPDPLFNPLKTVGCQDNNVNTDA-----ILSRAGGS--IADFTGHR 181
Db 141 PQALAT-----RALPCPAHVVFYQAGREDELITTEGEMLEVEEGDADSWKARNQHG 194
Qy 182 QTAFRELERVLPQSNNCLKREKQDESCS-----LTOALPSELKVSADNVSL-TGAVS 234
Db 195 EVGFVP-ERYLNFPLDLPSSQSDNPPCGAEPFTAFLAQALYSYTGQSAEELSFPFEGAL- 252
Qy 235 LASMLTIFLLQQAQGMPEPCWGRIITDHSQWNTLLSLHNAQFYLLQRTPEVARSRATPLL 294

Db 253 -----IRLLPRAQGVDDGFWRGEGRGVFPFSLVVEELLGPGPPE-----LS 297
Qy 295 DLIMAALTTPHPKQKQAYGVTLPTSVL 320
Db 298 DPEQMLPSPSP---SFSPPAPTSLV 320
RESULT 14
US-11-293-697-3258
; Sequence 3258, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3258
; LENGTH: 725
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-3258
Query Match 3.6%; Score 84; DB 7; Length 725;
Best Local Similarity 19.9%; Pred. No. 11;
Matches 68; Conservative 52; Mismatches 92; Indels 130; Gaps 19;
Qy 119 GEAPAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQDNNVNTDAILSRAGGSIAFT 178
Db 23 GQESTFGIVPD-AVQVHTERQQS-----HPSES---LD---TDSLL-----IAVSS 62
Qy 179 GHRQTAFRELERVLPQSNNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASM 238
Db 63 QH---TVNTELEKQIS-----NEVDSE-----DLKMSSEVKHICSEDQIEDK 101
Qy 239 LTBIFLLQQAQGMPEPCWGRIITD-----SHOWNTLLSLHNAQFYLLQRTPEVARSRATPL 293
Db 102 M-----EVTENIEVHTQ---ITVQEQQLLLEE-PETVVSR--- 135
Qy 294 LDIMAALTTPHPKQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNTLPGQDNTPPGG 353
Db 136 -----ESRPPKLVNESVTLPLETL-VSPHEESI-----SLCPPE 168
Qy 354 ELVFERWRRLSDNSQWTVQSLVFTLQQMRDKTPLSLNTPPGEVKLTLAGC-BERNAQGM 412
Db 169 QLVIER-----LQGEKEQENSELSLGLMDSEMTPTIEGCVKDVSYQGG 212
Qy 413 CSL-----AGTQI-----VNEARI PACSLRSHHHHH 440
Db 213 KSIKLSSETESFSSSADISKADVSSPTSPSDLPSPHMDLHN 254

RESULT 15

US-10-953-349-22123
; Sequence 22123, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DERIVED
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 22123
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Glycine max

Search completed: June 13, 2006, 10:51:02
Job time : 12.0917 secs

THIS PAGE LEFT BLANK

Qy 1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPWP 60
 Db 1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPWP 60
 Qy 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPSQGVALLIADVDERTRKTGE 120
 Db 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPSQGVALLIADVDERTRKTGE 120
 Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQLDANVTDAILSRAGGSIADFTGH 180
 Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQLDANVTDAILSRAGGSIADFTGH 180
 Qy 181 ROTAFRELERVLNFPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 Db 181 ROTAFRELERVLNFPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 Qy 241 EIFLLOQAQMPGPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLDLLIMAA 300
 Db 241 EIFLLOQAQMPGPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLDLLIMAA 300
 Qy 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLNLGGALELNWTLPGQPDNTPPGGSELVFERW 360
 Db 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLNLGGALELNWTLPGQPDNTPPGGSELVFERW 360
 Qy 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCMSLAGFTQ 420
 Db 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCMSLAGFTQ 420
 Qy 421 IVNEARIPACSLRSHHHHH 440
 Db 421 IVNEARIPACSLRSHHHHH 440

RESULT 2

AAE22836
 ID AAE22836 standard; protein; 440 AA.
 XX
 AC AAE22836;
 XX
 DT 21-AUG-2002 (first entry)
 XX
 DE Escherichia coli phytase protein.
 XX
 KW Dietary aid; biocompatible composition; therapeutic; digestive tract;
 KW foodstuff; digestion; phytase; enzyme.
 XX
 OS Escherichia coli.
 XX
 PN WO200189317-A2.
 XX
 PD 29-NOV-2001.
 XX
 PP 15-MAY-2001; 2001WO-US015764.
 XX
 PR 25-MAY-2000; 2000US-00580937.
 XX
 PA (DIVE-) DIVERSA CORP.
 XX
 PI Short JM, Kretz KA, O'donoghue E;
 XX
 XX WPI; 2002-164149/21.
 DR
 DR N-PSDB; AAD36473.

New dietary aids comprising sustained release biocompatible compositions, comprise agent that assists in digestion, useful for delivering enzymes, therapeutics, medicine or agents to an organism.
 XX
 PS Claim 5; Fig 1; 89pp; English.
 XX
 CC The present invention relates to novel dietary aids comprising sustained release biocompatible composition which comprises an agent (enzymes such as phytase, amylase, esterase, protease) that assists in digestion. The

CC biocompatible composition is effective upon oral consumption and release in the digestive tract of a subject. The dietary aids are useful for delivering enzymes, therapeutics, medicine and agents to an organism. The use of enzymes and other agents in digestive aids of livestock or domesticated animals not only improves the animal's health and life expectancy but also assists in increasing the health of livestock or in the production of foodstuffs from livestock. The present sequence is Escherichia coli phytase protein

XX
 SQ Sequence 440 AA;

Query Match 100.0%; Score 2302; DB 5; Length 440;
 Best Local Similarity 100.0%; Pred No. 2.7e-223;
 Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPWP 60
 Db 1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPWP 60
 Qy 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPSQGVALLIADVDERTRKTGE 120
 Db 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPSQGVALLIADVDERTRKTGE 120
 Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQLDANVTDAILSRAGGSIADFTGH 180
 Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQLDANVTDAILSRAGGSIADFTGH 180
 Qy 181 ROTAFRELERVLNFPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 Db 181 ROTAFRELERVLNFPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 Qy 241 EIFLLOQAQMPGPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLDLLIMAA 300
 Db 241 EIFLLOQAQMPGPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLDLLIMAA 300
 Qy 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLNLGGALELNWTLPGQPDNTPPGGSELVFERW 360
 Db 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLNLGGALELNWTLPGQPDNTPPGGSELVFERW 360
 Qy 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCMSLAGFTQ 420
 Db 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCMSLAGFTQ 420
 Qy 421 IVNEARIPACSLRSHHHHH 440
 Db 421 IVNEARIPACSLRSHHHHH 440

RESULT 3

AAE15806
 ID AAE15806 standard; protein; 440 AA.
 XX
 AC AAE15806;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Escherichia coli B phytase protein.
 XX
 KW Bacterial phytase; K12 appA phytase; protease stability; anabolic;
 KW gastrointestinal; nutritional value; feed treatment process; therapy;
 KW thermal tolerance; growth performance; alcoholic drink; biopulping;
 KW non-alcoholic drink; bioleaching; B phytase.
 XX
 OS Escherichia coli.
 XX
 PN WO200190333-A2.
 XX
 PD 29-NOV-2001.
 XX
 PP 24-MAY-2001; 2001WO-US017118.
 XX
 PR 25-MAY-2000; 2000US-00580515.

XX
 XX as phytase, amylase, esterase, protease) that assists in digestion. The

PA (DIVE-) DIVERSA CORP.
XX Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, O'donoghue E;
XX MPI; 2002-083108/11.
XX N-PSDB; AAD25460.
XX
XX New bacterial phytase for e.g. improving the nutritional value of phytate
XX -containing foodstuffs and subsequently improving the growth performance
XX of an organism that consumes it, or in treating animal digestive systems.
XX
XX Claim 1; Fig 1; 170pp; English.
XX
XX The patent discloses recombinant bacterial phytase from Escherichia coli
XX K12 appA phytase. The enzyme has phytase activity and improved thermal
XX tolerance when compared with wild-type phytase. It has improved protease
XX stability at low pH. The recombinant phytase is useful for improving the
XX nutritional value of phytate-containing foodstuffs and subsequently
XX improving the growth performance of an organism that consumes it, in
XX treating animal digestive systems, in feed treatment processes and for in
XX vitro purposes related to research, discovery and development. They are
XX also used for generating recombinant digestive system life forms, for
XX producing or manufacturing alcoholic and non-alcoholic drinks based on
XX the use of moulds, grains and/or plants, in biopulping and bio- bleaching
XX where a reduction in the use of environmentally harmful chemicals that
XX are traditionally used in the pulp and paper industry is desired and in
XX the reduction or possible elimination of the need for mineral
XX supplements, enzymes or therapeutic drugs for animals from the daily feed
XX thus increasing the amount calories and nutrients present in the feed.
XX The present sequence is E. coli B phytase protein
XX
SQ Sequence 440 AA;

Query Match 100.0%; Score 2302; DB 5; Length 440;
Best Local Similarity 100.0%; Pred. No. 2.7e-223;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKAILPFLSLILPTQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPAMP 60
Db 1 MKAILPFLSLILPTQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPAMP 60

Qy 61 TWPVKLWLTGRGELIAYLGHYQRLVADGLLAKGCPQSGQVAILADVDERTRKTGE 120
Db 61 TWPVKLWLTGRGELIAYLGHYQRLVADGLLAKGCPQSGQVAILADVDERTRKTGE 120

Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLENPLKTVGCOLDNANVTDAILSRAGGSTADFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLENPLKTVGCOLDNANVTDAILSRAGGSTADFTGH 180

Qy 181 RQTAFLERLVNFPQSNCLKREKQDESCSLTQALPSELKVSADNVSILGAVSLASMLT 240
Db 181 RQTAFLERLVNFPQSNCLKREKQDESCSLTQALPSELKVSADNVSILGAVSLASMLT 240

Qy 241 EIFLLOAQGPMPGCGWGRITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLDLIMAA 300
Db 241 EIFLLOAQGPMPGCGWGRITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLDLIMAA 300

Qy 301 LTPHPPOKQVGVLTPTSVLFIAGHDNLANLGALBNLWTLPGQDNTPPGSELVFERW 360
Db 301 LTPHPPOKQVGVLTPTSVLFIAGHDNLANLGALBNLWTLPGQDNTPPGSELVFERW 360

Qy 361 RRLSDNSQWQVSLVFTQLQOMRDKTPLSLNTPPGVEKVLTLAGEEENNAQWCSLAGFTQ 420
Db 361 RRLSDNSQWQVSLVFTQLQOMRDKTPLSLNTPPGVEKVLTLAGEEENNAQWCSLAGFTQ 420

Qy 421 IVNEARIPACSLRSHHHHHH 440
Db 421 IVNEARIPACSLRSHHHHHH 440

RESULT 4
ADA19446
ID ADA19446 standard; protein; 440 AA.

XX ADA19446;
XX
XX 20-NOV-2003 (first entry)
XX
XX E. coli B phytase.
XX
XX Phytase; enzyme; phytate; appA gene; animal feed; inorganic phosphate;
XX digestion enhancement; transgenic; thermal tolerance; protease stability.
XX
XX Escherichia coli B.
XX
XX Key Location/Qualifiers
XX Misc-difference 72 /note= "Encoded by CGN"
XX
XX US2002136754-A1.
XX
XX 26-SEP-2002.
XX
XX 24-MAY-2001; 2001US-00866379.
XX
XX 13-AUG-1997; 97US-00910798.
XX 01-MAR-1999; 99US-00259214.
XX 13-APR-1999; 99US-00291931.
XX 25-MAY-1999; 99US-00318528.
XX 25-MAY-2000; 2000US-00580515.
XX
XX (SHOR/) SHORT J M.
XX (KRET/) KRETZ K A.
XX (GRAY/) GRAY K A.
XX (BART/) BARTON N R.
XX (GARR/) GARRETT J B.
XX (DONO/) O' DONOGHUE E.
XX (MATH/) MATHUR E J.
XX
XX Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB;
XX O' Donoghue E, Mathur EJ;
XX MPI; 2003-040002/03.
XX
XX Isolated Escherichia coli polynucleotide encoding a modified phytase
XX enzyme, useful in the production of animal feed, for improving the
XX nutritional value of phytate-containing foodstuff and for enhancing
XX digestion in humans and animals.
XX
XX Claim 1; Fig 1; 62pp; English.
XX
XX The invention relates to an isolated Escherichia coli polynucleotide
XX encoding a phytase enzyme appearing as ADA19450 and having amino acids
XX modification from W68E, Q84W, A95P, K97C, S168E, R180Y, N226C or Y277D.
XX Also included the E. coli appA gene ADA19449 (or an oligonucleotide
XX derived from it) or its mutant sequence ADA19452, expression vectors,
XX host cells, a method of improving nutritional value of a phytate-
XX containing foodstuff by contacting the phytate-containing foodstuff with
XX a pure phytase enzyme mutated phytase (where the phytase enzyme catalyzes
XX the liberation of inorganic phosphate from the phytate in the phytate-
XX containing foodstuff), a method to produce an animal feed containing a
XX microbial phytase (comprising culturing the plant cell, plant part or
XX plant under conditions where the nucleotide sequence is expressed and
XX converting the plant cells, plants or plant into a composition for animal
XX feed), a feed composition for animals (comprising the plant seeds, plant
XX cells, plant parts or plants in admixture with a phytate-containing
XX foodstuff), a method to treat a human or an animal able to benefit from
XX digestive enhancement by the activity of an exogenous phytase enzymes
XX comprising administering to the human or animal the plant seed, plant
XX cells, plant parts or plants of a transgenic plant which is modified to
XX contain an expression system which expresses a nucleotide sequence
XX encoding a phytase enzyme, a transgenic non-human organism whose genome
XX comprising a heterologous nucleic acid sequence encoding a polypeptide
XX having phytase activity. The phytase enzyme is useful for improving the
XX nutritional value of phytate-containing foodstuff, in the production of
XX animal feed and for enhancing digestion in humans and animals. The

CC invented method improves thermal tolerance and protease stability. It
CC also improves the feeding value of phytate rich ingredients. The present
CC sequence represents E. coli B wild-type phytase.
XX
SQ Sequence 440 AA;

Query Match 100.0%; Score 2302; DB 6; Length 440;
Best Local Similarity 100.0%; Pred. No. 2.7e-223;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
Db 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
Qy 61 TWPVKLGWLTFRGGELIAYLGHYQORLVADGLLAKKGCPSQGVAILIADVDRTRKTGE 120
Db 61 TWPVKLGWLTFRGGELIAYLGHYQORLVADGLLAKKGCPSQGVAILIADVDRTRKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIADFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIADFTGH 180
Qy 181 RQAFRELERVLNPPQSNLCIKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 RQAFRELERVLNPPQSNLCIKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLLOAQGMPEFGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSATPLLDLIMAA 300
Db 241 EIFLLOAQGMPEFGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSATPLLDLIMAA 300
Qy 301 LTPHPKQKAYGVTLPTSLVFIAGHDTNLNLANLGALNWLTPQPDNTPPGGELVPERW 360
Db 301 LTPHPKQKAYGVTLPTSLVFIAGHDTNLNLANLGALNWLTPQPDNTPPGGELVPERW 360
Qy 361 RRLSDNSQWIQVSLVFTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Db 361 RRLSDNSQWIQVSLVFTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Qy 421 IVNEARIPACSLRSHHHHH 440
Db 421 IVNEARIPACSLRSHHHHH 440

RESULT 5
ADO50292
ID ADO50292 standard; protein; 440 AA.
AC ADO50292;
DT 29-JUL-2004 (first entry)
DE Escherichia coli B phytase enzyme.
XX
KW Phytase; bacteria; thermal tolerance; protease stability; foodstuff;
KW phytate; animal feed; fish feed; dough; baking; enzyme.
XX
OS Escherichia coli; B.
XX
FH Key Location/Qualifiers
FT Misc-difference 72
FT /note= "Encoded by CGN"
XX
PN US2004091968-A1.
XX
PD 13-MAY-2004.
XX
XX 20-JUN-2003; 2003US-00601319.
XX
XX 13-AUG-1997; 97US-00910798.
PR 01-MAR-1999; 99US-00259214.
PR 13-APR-1999; 99US-00291931.
PR 25-MAY-1999; 99US-00318528.
PR 25-MAY-2000; 2000US-00580515.

PR
XX
PA (SHOR/) SHORT J M.
PA (KRET/) KRETZ K.
PA (GRAY/) GRAY K A.
PA (BART/) BARTON N R.
PA (GARE/) GARRETT J B.
PA (ODON/) O'DONOGHUE E.
PA (MATH/) MATHER E J.
XX
PI Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI Mather EJ;
DR WPI; 2004-374952/35.
XX N-PSDB; ADO50291.
XX
PT Producing phytase, involves providing nucleic acid derived from bacteria
PT encoding polypeptide having phytase activity, and expressing nucleic acid
PT in yeast.
XX
PS Claim 3; SEQ ID NO 2; 74pp; English.
XX
CC The invention relates to a method for producing phytase that involves
CC providing a nucleic acid encoding phytase derived from a bacteria, and
CC expressing the nucleic acid in a yeast under conditions that allow
CC expression of the enzyme in the yeast. The invention also relates to
CC modified phytase enzyme which has improved thermal tolerance and protease
CC stability at low pH. The phytase enzyme can be used in foodstuffs to
CC improve the feeding value of phytate rich ingredients, and in diet of
CC numerous animals including mammals, fowls and fishes, commercially
CC significant mammals such as pigs, goats, laboratory rodents, commercially
CC significant avian species such as chicken, ducks, doves, parrot, etc.,
CC commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC etc., in dough making and baking, in dietary aids for animals. The method
CC provides easy manufacture of the active ingredient loaded biocompatible
CC composition, higher yields and loading efficiency. The phytase
CC incorporated in to the dietary aid is safe for animals. The present
CC sequence is Escherichia coli B phytase enzyme.

SQ Sequence 440 AA;

Query Match 100.0%; Score 2302; DB 8; Length 440;

Best Local Similarity 100.0%; Pred. No. 2.7e-223;

Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQMLQMDVTPDAMP 60

Db 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQMLQMDVTPDAMP 60

Qy 61 TWPVKLGWLTFRGGELIAYLGHYQORLVADGLLAKKGCPSQGVAILIADVDRTRKTGE 120

Db 61 TWPVKLGWLTFRGGELIAYLGHYQORLVADGLLAKKGCPSQGVAILIADVDRTRKTGE 120

Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIADFTGH 180

Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIADFTGH 180

Qy 181 RQAFRELERVLNPPQSNLCIKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Db 181 RQAFRELERVLNPPQSNLCIKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy 241 EIFLLOAQGMPEFGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSATPLLDLIMAA 300

Db 241 EIFLLOAQGMPEFGWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSATPLLDLIMAA 300

Qy 301 LTPHPKQKAYGVTLPTSLVFIAGHDTNLNLANLGALNWLTPQPDNTPPGGELVPERW 360

Db 301 LTPHPKQKAYGVTLPTSLVFIAGHDTNLNLANLGALNWLTPQPDNTPPGGELVPERW 360

Qy 361 RRLSDNSQWIQVSLVFTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Db 361 RRLSDNSQWIQVSLVFTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

QY 421 IVNEARIPACSLRSHHHHH 440
 Db |||||

RESULT 6

AE75413
 ID AEE75413 standard; protein; 440 AA.

XX AC AEE75413;

XX DT 23-FEB-2006 (first entry)

XX DE E. coli B phytase SEQ ID NO: 2.

XX KW enzyme; thermostable; phytase; protein stabilization; pharmaceutical;
 XX fertilizier; cramp; cell culture; osteoporosis; osteopathic; food.

XX OS Escherichia coli B.

XX FH Key Location/Qualifiers

XX FT Misc-difference 72 /note= "Encoded by CGN"

XX FN US2005281792-A1.

XX PD 22-DEC-2005.

XX PF 01-SEP-2004; 2004US-00933115.

XX PR 13-AUG-1997; 97US-00910798.

XX PR 01-MAR-1999; 99US-00259214.

XX PR 13-APR-1999; 99US-00291931.

XX PR 25-MAY-1999; 99US-00318528.

XX PR 25-MAY-2000; 2000US-00580515.

XX PR 24-MAY-2001; 2001US-00866379.

XX PA (SHOR/) SHORT J M.

XX PA (KRET/) KRETZ K A.

XX PA (GRAY/) GRAY K A.

XX PA (BART/) BARTON N R.

XX PA (GARR/) GARRETT J B.

XX PA (ODON/) O'DONOGHUE E.

XX PA (BAUM/) BAUM W.

XX PA (ROBE/) ROBERTSON D E.

XX PA (ZORN/) ZORNER P.

XX PI Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, Odonoghue E;
 PI Baum W, Robertson DE, Zorner P;

XX WPI; 2006-055022/06.

XX N-PSDB; AEE75412.

XX FT Formulation useful as dietary supplement for treating, preventing or
 XX reversing osteoporosis and bone loss, and preventing muscle cramps,
 XX comprises one or more polypeptides having phytase activity.

XX PS Claim 1; SEQ ID NO 2; 82pp; English.

XX CC The present sequence is that of the *Escherichia coli* B phytase. The
 CC present invention relates to a novel formulation, comprising at least one
 CC polypeptide having phytase activity, where the polypeptide is a modified
 CC sequence (AEE75421) derived from the *Escherichia coli* K-12 appA phytase
 CC (AEE75419). The modification of the enzyme, comprising a series of
 CC mutations, improves the thermal tolerance and protease stability of the
 CC protein. The specification also claims a pharmaceutical composition; a
 CC kit; an immobilized phytase; a fertilizer or soil additive; a liquid
 CC supplement for preventing muscle cramps; a hydrating agent; a tissue
 CC culture or cell culture media; and a plant food additive, all comprising
 CC the novel phytase, and a method of reducing pollution and increasing
 CC nutrient availability in an environment or environmental sample by
 CC degrading environmental phytic acid. The novel enzyme is a dietary
 CC supplement useful for treating, preventing or reversing osteoporosis or

CC bone loss, and preventing muscle cramps. The liquid supplement is useful
 CC for preventing muscle cramps. The method is useful for reducing pollution
 CC and increasing nutrient availability in an environment or environmental
 CC sample by degrading environmental phytic acid, where the environment or
 CC environmental sample comprises a soil or a body of water. The immobilized
 CC phytase is useful in foodstuffs for improving the feeding value of
 CC phytate rich ingredients.

XX SQ Sequence 440 AA;

Query Match 100.0%; Score 2302; DB 10; Length 440;

Best Local Similarity 100.0%; Pred. No. 2.7e-223;

Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAILIPLSLLIPTQSAFAQSEPELKSWSVIVSRHGVRAPTKATQLMDVTPDAMP 60

Db 1 MKAILIPLSLLIPTQSAFAQSEPELKSWSVIVSRHGVRAPTKATQLMDVTPDAMP 60

QY 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGVVAIIADVDERTKTGE 120

Db 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGVVAIIADVDERTKTGE 120

QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAADFTGH 180

Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAADFTGH 180

QY 181 ROTAPRELRLVNFQSNLCIKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Db 181 ROTAPRELRLVNFQSNLCIKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

QY 241 EIFLLQAQGMPEPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLDLIMAA 300

Db 241 EIFLLQAQGMPEPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLDLIMAA 300

QY 301 LTPHPQKQAYGVTLPTSVLFIAGHDVNLNGLGALNLNWTLPQDPNTPPGGELVFERW 360

Db 301 LTPHPQKQAYGVTLPTSVLFIAGHDVNLNGLGALNLNWTLPQDPNTPPGGELVFERW 360

QY 361 RRLSDNSQWIOVSLVFTLQOMRDKTLPISLNTTPGVEVKLTLAGCERNAGMCSLAGFTQ 420

Db 361 RRLSDNSQWIOVSLVFTLQOMRDKTLPISLNTTPGVEVKLTLAGCERNAGMCSLAGFTQ 420

QY 421 IVNEARIPACSLRSHHHHH 440

Db 421 IVNEARIPACSLRSHHHHH 440

RESULT 7

AD050304

ID AD050304 standard; protein; 440 AA.

XX AC AD050304;

XX DT 29-JUL-2004 (first entry)

XX DE *Escherichia coli* W phytase 875PH2 mutant enzyme.

XX KW Phytase; bacteria; thermal tolerance; protease stability; foodstuffs;
 XX phytate; animal feed; fish feed; dough; baking; mutant; mutain.

XX OS *Escherichia coli*.

XX FH Key Location/Qualifiers

XX FT Misc-difference 72 /note= "Encoded by CGN"

XX FT Misc-difference 160 /note= "E. coli B phytase Ala replaced with Ser"

XX FT Misc-difference 176 /note= "E. coli B phytase Asp replaced with Gly"

XX FT Misc-difference 298 /note= "E. coli B phytase Met replaced with Lys"

XX FT Misc-difference 299 /note= "E. coli B phytase Ala replaced with Thr"

XX PN US2004091968-A1.
XX XX 13-MAY-2004.
XX PF 20-JUN-2003; 2003US-00601319.
XX PR 13-AUG-1997; 97US-00910798.
XX PR 01-MAR-1999; 99US-00259214.
XX PR 13-APR-1999; 99US-00291931.
XX PR 25-MAY-1999; 99US-00318528.
XX PR 25-MAY-2000; 2000US-00580515.
XX PR 24-MAY-2001; 2001US-00866379.
XX (SHOR/) SHORT J M.
XX PA (KRET/) KRETZ K.
XX PA (GRAY/) GRAY K A.
XX PA (BART/) BARTON N R.
XX PA (GARR/) GARRETT J B.
XX PA (ODON/) O'DONOGHUE E.
XX PA (MATH/) MATHER E J.
XX Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
XX Mather EJ;
XX
XX WPI; 2004-374952/35.
XX DR N-PSDB; ADO50291, ADO50303.
XX
XX Producing phytase, involves providing nucleic acid derived from bacteria
XX encoding polypeptide having phytase activity, and expressing nucleic acid
XX in yeast.
XX
XX Disclosure; Page; 74pp; English.
XX
XX The invention relates to a method for producing phytase that involves
XX providing a nucleic acid encoding phytase derived from a bacteria, and
XX expressing the nucleic acid in a yeast under conditions that allow
XX expression of the enzyme in the yeast. The invention also relates to
XX modified phytase enzyme which has improved thermal tolerance and protease
XX stability at low pH. The phytase enzyme can be used in foodstuffs to
XX improve the feeding value of phytate rich ingredients, and in diet of
XX numerous animals including mammals, fowls and fishes, commercially
XX significant mammals such as pigs, goats, laboratory rodents, commercially
XX significant avian species such as chicken, ducks, doves, parrot, etc.,
XX commercially farmed fish such as guppy, zebrafish, molly, swordtail,
XX etc., in dough making and baking, in dietary aids for animals. The method
XX provides easy manufacture of the active ingredient loaded biocompatible
XX composition, higher yields and loading efficiency. The phytase
XX incorporated in to the dietary aid is safe for animals. The present
XX sequence is Escherichia coli w phytase mutant enzyme. Note: The present
XX sequence is not shown in the specification but has been derived from
XX Escherichia coli B phytase ADO50292.
XX
XX Sequence 440 AA;
XX
XX Query Match 99.1%; Score 2282; DB 8; Length 440;
XX Best Local Similarity 99.1%; Pred. No. 2.9e-221;
XX Matches 436; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
XX
XX 1 MKAILIPFLSLLIPTQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPAMP 60
XX
XX 1 MKAILIPFLSLLIPTQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPAMP 60
XX
XX 61 TPVPKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
XX
XX 61 TPVPKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
XX
XX 121 AFAAGLAPDCNITVHTQADTSSPPDLFNPLKTGVCOLDNANVTAILSRAGGSTADFTGH 180
XX
XX 121 AFAAGLAPDCNITVHTQADTSSPPDLFNPLKTGVCOLDNANVTAILSRAGGSTAGFTGH 180
XX
XX 181 RQTAFRELERVLPQSNLCIKRKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
XX

Db 181 RQTAFRELERVLPQSNLCIKRKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFFLQAQGMPEPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Db 241 EIFFLQAQGMPEPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300
Qy 301 LTPHPQKQAYGVTLPTSULFIAGHDNTLANLGGALELNWTLPGQPDNTPGGELVFERW 360
Db 301 LTPHPQKQAYGVTLPTSULFIAGHDNTLANLGGALELNWTLPGQPDNTPGGELVFERW 360
Qy 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTTPGGEVKLTLAGCEERNAQCMCSLAGFTQ 420
Db 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTTPGGEVKLTLAGCEERNAQCMCSLAGFTQ 420
Qy 421 INVNEARIPACSLRSHHHHH 440
Db 421 INVNEARIPACSLRSHHHHH 440
RESULT 8
ADO50302
ID ADO50302 standard; protein; 440 AA.
XX AC ADO50302;
XX DT 29-JUL-2004 (first entry)
XX DE Kangaroo rat Escherichia coli phytase 872PH1 mutant enzyme.
XX KW Phytase; bacteria; thermal tolerance; protease stability; foodstuff;
XX KW phytate; animal feed; fish feed; dough; baking; mutant; mutein.
XX OS Escherichia coli.
XX Key Location/Qualifiers
XX Misc-difference 72 /note= "Encoded by CGN"
XX FT Misc-difference 176 /note= "E. coli B phytase Asp replaced with Gly"
XX FT Misc-difference 298 /note= "E. coli B phytase Met replaced with Lys"
XX FT Misc-difference 299 /note= "E. coli B phytase Ala replaced with Thr"
XX FT Misc-difference 312 /note= "E. coli B phytase Gly replaced with Ser"
XX US2004091968-A1.
XX PD 13-MAY-2004.
XX PF 20-JUN-2003; 2003US-00601319.
XX PR 13-AUG-1997; 97US-00910798.
XX PR 01-MAR-1999; 99US-00259214.
XX PR 13-APR-1999; 99US-00291931.
XX PR 25-MAY-1999; 99US-00318528.
XX PR 25-MAY-2000; 2000US-00580515.
XX PR 24-MAY-2001; 2001US-00866379.
XX (SHOR/) SHORT J M.
XX PA (KRET/) KRETZ K.
XX PA (GRAY/) GRAY K A.
XX PA (BART/) BARTON N R.
XX PA (GARR/) GARRETT J B.
XX PA (ODON/) O'DONOGHUE E.
XX PA (MATH/) MATHER E J.
XX Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
XX Mather EJ;
XX WPI; 2004-374952/35.
XX DR N-PSDB; ADO50301.
XX

PT Producing phytase, involves providing nucleic acid derived from bacteria
 PT encoding polypeptide having phytase activity, and expressing nucleic acid
 XX in yeast.
 PS Disclosure; Page; 74pp; English.
 XX
 XX The invention relates to a method for producing phytase that involves
 CC providing a nucleic acid encoding phytase derived from a bacteria, and
 CC expressing the nucleic acid in a yeast under conditions that allow
 CC expression of the enzyme in the yeast. The invention also relates to
 CC modified phytase enzyme which has improved thermal tolerance and protease
 CC stability at low pH. The phytase enzyme can be used in feedstuffs to
 CC improve the feeding value of phytate rich ingredients, and in diet of
 CC numerous animals including mammals, fowls and fishes, commercially
 CC significant mammals such as pigs, goats, laboratory rodents, commercially
 CC significant avian species such as chicken, ducks, doves, parrot, etc.,
 CC commercially farmed fish such as guppy, zebrafish, molly, swordtail,
 CC etc., in dough making and baking, in dietary aids for animals. The method
 CC provides easy manufacture of the active ingredient loaded biocompatible
 CC composition, higher yields and loading efficiency. The phytase
 CC incorporated in to the dietary aid is safe for animals. The present
 CC sequence is kangaroo rat *Escherichia coli* phytase mutant enzyme. Note:
 CC The present sequence is not shown in the specification but has been
 CC derived from *Escherichia coli* B phytase sequence ADO50292.
 XX

Sequence 440 AA;

Query Match 99.0%; Score 2279; DB 8; Length 440;
 Best Local Similarity 99.1%; Pred. No. 5.7e-221;
 Matches 436; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 MKAILPFLSLIPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
 DB 1 MKAILPFLSLIPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
 QY 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPOSQGVVAIADVDERTRKTGE 120
 DB 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPOSQGVVAIADVDERTRKTGE 120
 QY 121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGSIADFTGH 180
 DB 121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGSIADFTGH 180
 QY 181 RQTAPRELRLVNPQSNLCLKREKQDESCSLTQALPSELKVSADNVSILTGAVSLASMLT 240
 DB 181 RQTAPRELRLVNPQSNLCLKREKQDESCSLTQALPSELKVSADNVSILTGAVSLASMLT 240
 QY 241 EIFLLQQAQMPGPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLINKTA 300
 DB 241 EIFLLQQAQMPGPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLINKTA 300
 QY 301 LTPHPQQAQYGVTLPTSVLFIAGHDNTNLNAGALELNWTLPGQPDNTPPGGELVFERW 360
 DB 301 LTPHPQQAQYGVTLPTSVLFIAGHDNTNLNAGALELNWTLPGQPDNTPPGGELVFERW 360
 QY 361 RLSDNSQWIOVSLVFQTLQMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCSLAGFTQ 420
 DB 361 RLSDNSQWIOVSLVFQTLQMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCSLAGFTQ 420
 QY 421 IVNEARIPACSLRSHHHHH 440
 DB 421 IVNEARIPACSLRSHHHHH 440

RESULT 9

AAB36257
 ID AAB36257 standard; protein; 432 AA.
 XX
 XX
 AC AAB36257;
 XX
 XX
 DT 12-SEP-2003 (revised)
 XX 20-FEB-2001 (first entry)
 XX

DE Lama2/APPA plasmid translated sequence.
 XX Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
 KW environmental pollution; pig.
 XX
 XX *Mus musculus*.
 OS *Escherichia coli*.
 OS *Chimeric*.
 XX
 XX WO200064247-A1.
 PN
 XX
 XX 02-NOV-2000.
 PD
 XX
 XX 20-APR-2000; 2000WO-CA000430.
 PF
 XX
 XX 23-APR-1999; 99US-0130508P.
 PR
 XX (UYGU-) UNIV GUELPH.
 PA
 XX Forsberg CW, Golovan S, Phillips JP;
 PI
 XX WPI; 2000-687245/67.
 DR
 XX N-PSDB; AAC68294.
 DR
 XX Transgenic non-human animal for gastrointestinal tract specific
 PT expression of a protein, preferably phytase, comprises a nucleic acid
 PT sequence including a heterologous transgene construct encoding the
 PT protein.
 PT
 XX Disclosure; Fig 5; 152pp; English.
 PS
 XX The present invention provides transgenic animals which produce desired
 CC proteins, in this case pigs which expresses phytase in the salivary
 CC gland. Low phytase production levels result in phytate in the diet being
 CC excreted and causing phosphorus contamination in water, as well as
 CC reducing the growth of animals. The invention provides a number of
 CC transgenes containing the *E. coli* APPA phytase coding sequence. (Updated
 CC on 12-SEP-2003 to standardise OS field)
 XX
 SQ Sequence 432 AA;

Query Match 97.1%; Score 2235; DB 3; Length 432;
 Best Local Similarity 99.5%; Pred. No. 1.6e-216;
 Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MKAILPFLSLIPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
 DB 1 MKAILPFLSLIPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
 QY 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPOSQGVVAIADVDERTRKTGE 120
 DB 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPOSQGVVAIADVDERTRKTGE 120
 QY 121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGSIADFTGH 180
 DB 121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGSIADFTGH 180
 QY 181 RQTAPRELRLVNPQSNLCLKREKQDESCSLTQALPSELKVSADNVSILTGAVSLASMLT 240
 DB 181 RQTAPRELRLVNPQSNLCLKREKQDESCSLTQALPSELKVSADNVSILTGAVSLASMLT 240
 QY 241 EIFLLQQAQMPGPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLINKTA 300
 DB 241 EIFLLQQAQMPGPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLINKTA 300
 QY 301 LTPHPQQAQYGVTLPTSVLFIAGHDNTNLNAGALELNWTLPGQPDNTPPGGELVFERW 360
 DB 301 LTPHPQQAQYGVTLPTSVLFIAGHDNTNLNAGALELNWTLPGQPDNTPPGGELVFERW 360
 QY 361 RLSDNSQWIOVSLVFQTLQMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCSLAGFTQ 420
 DB 361 RLSDNSQWIOVSLVFQTLQMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCSLAGFTQ 420

```

QY 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432

RESULT 10
ID AAB36261 standard; protein; 432 AA.
XX AAB36261;
AC AAB36261;
XX
XX 12-SEP-2003 (revised)
DT 20-FEB-2001 (first entry)
XX
DE R15/APPA plasmid translated sequence.
XX
XX Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KW environmental pollution; pig.
XX
XX Rattus sp.
OS Escherichia coli.
OS Chimeric.
XX
XX WO200064247-A1.
PN
XX
XX 02-NOV-2000.
PD
XX
XX 20-APR-2000; 2000WO-CA000430.
PF
XX
XX 23-APR-1999; 99US-0130508P.
PR
XX
XX (UYGU-) UNIV GUELPH.
PA
XX
XX Forsberg CW, Golovan S, Phillips JP;
PI
XX WPI; 2000-687245/67.
DR
XX N-PSDB; AAC68298.
DR
XX
XX Transgenic non-human animal for gastrointestinal tract specific
PT expression of a protein, preferably phytase, comprises a nucleic acid
PT sequence including a heterologous transgene construct encoding the
PT protein.
XX
XX Disclosure; Fig 21; 152pp; English.
XX
XX The present invention provides transgenic animals which produce desired
CC proteins, in this case pigs which expresses phytase in the salivary
CC gland. Low phytase production levels result in phytate in the diet being
CC excreted and causing phosphorus contamination in water, as well as
CC reducing the growth of animals. The invention provides a number of
CC transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC on 12-SEP-2003 to standardise OS field)
XX
XX Sequence 432 AA;
SQ
Query Match 97.1%; Score 2235; DB 3; Length 432;
Best Local Similarity 99.5%; Pred. No. 1.6e-216;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKAILIFLSLLIPTQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPWP 60
Db 1 MKAILIFLSLLIPTQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPWP 60

QY 61 TWPVKLGLTPRGELIAYLGHYQRLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
Db 61 TWPVKLGLTPRGELIAYLGHYQRLVADGLLAKKGCPSQGVAIADVDERTKTGE 120

QY 121 AFAAGLAPDCAITVHTQADTSSPDPLENPLKTGVCOLDNANVTDAILSRAGGSITADFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLENPLKTGVCOLDNANVTDAILSRAGGSITADFTGH 180

QY 181 RQTAFRELERVLNFPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 RQTAFRELERVLNFPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

```

QY 1 MKAIIIPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
 |||||
 Db 1 MKAIIIPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
 |||||
 QY 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPOSQGVAAIADVDERTKTGE 120
 |||||
 Db 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPOSQGVAAIADVDERTKTGE 120
 |||||
 QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKGTGVCOLDNANVTDAILSRAGGSIAADFTGH 180
 |||||
 Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKGTGVCOLDNANVTDAILSRAGGSIAADFTGH 180
 |||||
 QY 181 ROTAFRELERVLNFPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 |||||
 Db 181 ROTAFRELERVLNFPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 |||||
 QY 241 EIFLLQQAQMGPEPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
 |||||
 Db 241 EIFLLQQAQMGPEPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300
 |||||
 QY 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
 |||||
 Db 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
 |||||
 QY 361 RRLSDNSQWIOVSLVFQTLQQRDKTPISLNTPPGEVKLTLAGCERNAQGCMSLAGFTQ 420
 |||||
 Db 361 RRLSDNSQWIOVSLVFQTLQQRDKTPISLNTPPGEVKLTLAGCERNAQGCMSLAGFTQ 420
 |||||
 QY 421 IVNEARIPACSL 432
 |||||
 Db 421 IVNEARIPACSL 432
 |||||

RESULT 12

AAB36259
 ID AAB36259 standard; protein; 432 AA.
 XX AC AAB36259;
 XX AC AAB36259;
 DT 12-SEP-2003 (revised)
 DT 20-FEB-2001 (first entry)
 XX XX
 DE R15/APPA plasmid translated sequence.
 XX Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
 KW environmental pollution; pig.
 XX Rattus sp.
 OS Escherichia coli.
 OS Chimeric.
 XX WO200064247-A1.
 XX 02-NOV-2000.
 XX 20-APR-2000; 2000WO-CA000430.
 XX 23-APR-1999; 99US-0130508P.
 XX (UYGU-) UNIV GUELPH.
 XX Forsberg CW, Golovan S, Phillips JP;
 XX WPI; 2000-687245/67.
 DR N-PSDB; AAC68296.
 XX Transgenic non-human animal for gastrointestinal tract specific
 PT expression of a protein, preferably phytase, comprises a nucleic acid
 PT sequence including a heterologous transgene construct encoding the
 PT protein.
 XX Disclosure; Fig 19; 152pp; English.

CC The present invention provides transgenic animals which produce desired
 CC proteins, in this case pigs which expresses phytase in the salivary
 CC gland. Low phytase production levels result in phytate in the diet being
 CC excreted and causing phosphorus contamination in water, as well as
 CC reducing the growth of animals. The invention provides a number of
 CC transgenes containing the E. coli APPA phytase coding sequence. (Updated
 CC on 12-SEP-2003 to standardise OS field)

SQ Sequence 432 AA;

Query Match 97.1%; Score 2235; DB 3; Length 432;
 Best Local Similarity 99.5%; Pred. No. 1.6e-216;
 Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKAIIIPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
 |||||
 Db 1 MKAIIIPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
 |||||
 QY 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPOSQGVAAIADVDERTKTGE 120
 |||||
 Db 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPOSQGVAAIADVDERTKTGE 120
 |||||
 QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKGTGVCOLDNANVTDAILSRAGGSIAADFTGH 180
 |||||
 Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKGTGVCOLDNANVTDAILSRAGGSIAADFTGH 180
 |||||
 QY 181 ROTAFRELERVLNFPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 |||||
 Db 181 ROTAFRELERVLNFPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 |||||
 QY 241 EIFLLQQAQMGPEPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
 |||||
 Db 241 EIFLLQQAQMGPEPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300
 |||||
 QY 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
 |||||
 Db 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
 |||||
 QY 361 RRLSDNSQWIOVSLVFQTLQQRDKTPISLNTPPGEVKLTLAGCERNAQGCMSLAGFTQ 420
 |||||
 Db 361 RRLSDNSQWIOVSLVFQTLQQRDKTPISLNTPPGEVKLTLAGCERNAQGCMSLAGFTQ 420
 |||||
 QY 421 IVNEARIPACSL 432
 |||||
 Db 421 IVNEARIPACSL 432
 |||||

RESULT 13

AAB36258
 ID AAB36258 standard; protein; 432 AA.
 XX AC AAB36258;
 XX AC AAB36258;
 DT 12-SEP-2003 (revised)
 DT 20-FEB-2001 (first entry)
 XX XX
 DE R15/APPA plasmid translated sequence.
 XX Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
 KW environmental pollution; pig.
 XX Rattus sp.
 OS Escherichia coli.
 OS Chimeric.
 XX WO200064247-A1.
 XX 02-NOV-2000.
 XX 20-APR-2000; 2000WO-CA000430.
 XX 23-APR-1999; 99US-0130508P.
 XX XX

PA (UYGU-) UNIV GUELPH.
XX Forsberg CW, Golovan S, Phillips JP;
XX WPI; 2000-687245/67.
XX N-PSDB; AAC68295.
XX Transgenic non-human animal for gastrointestinal tract specific
PT expression of a protein, preferably phytase, comprises a nucleic acid
PT sequence including a heterologous transgene construct encoding the
PT protein.
XX
XX
XX Disclosure; Fig 18; 152pp; English.
XX
XX The present invention provides transgenic animals which produce desired
CC proteins, in this case pigs which expresses phytase in the salivary
CC gland. Low phytase production levels result in phytate in the diet being
CC excreted and causing phosphorus contamination in water, as well as
CC reducing the growth of animals. The invention provides a number of
CC transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC on 12-SEP-2003 to standardise OS field)
XX
SQ Sequence 432 AA;
Query Match 97.1%; Score 2235; DB 3; Length 432;
Best Local Similarity 99.5%; Pred. No. 1.6e-216;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MKAILPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDWP 60
Db 1 MKAILPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDWP 60
Qy 61 TWPVKLGLTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGOVAIIADVDERTRKTGE 120
Db 61 TWPVKLGLTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGOVAIIADVDERTRKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAIILSRAGGSIA DPTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAIILSRAGGSIA DPTGH 180
Qy 181 RQTAFLRELRLVLPFQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 RQTAFLRELRLVLPFQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLLQQAQMPGPGWGRITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Db 241 EIFLLQQAQMPGPGWGRITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Qy 301 LTPHPKQAYGVTLPTSVLFIAGHDTNLNLGGALELNWTLPGQPDNTPPGGELVPERW 360
Db 301 LTPHPKQAYGVTLPTSVLFIAGHDTNLNLGGALELNWTLPGQPDNTPPGGELVPERW 360
Qy 361 RRLSDNSQWIVQSVLFTLQMRDKTPLSLNTPPEVKLTLAGCEERNAQCMCSLAGFTQ 420
Db 361 RRLSDNSQWIVQSVLFTLQMRDKTPLSLNTPPEVKLTLAGCEERNAQCMCSLAGFTQ 420
Qy 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432
RESULT 14
AAB36263
ID AAB36263 standard; protein; 432 AA.
XX AC AAB36263;
XX
XX
DT 12-SEP-2003 (revised)
DT 20-FEB-2001 (first entry)
XX
XX Lama2/APPA plasmid translated sequence.
XX Transgenic animal; salivary protein; phytase; phosphorus; animal growth;

KW environmental pollution; pig.
XX Mus musculus.
OS Escherichia coli.
XX Chimeric.
XX
XX W0200064247-A1.
XX
XX 02-NOV-2000.
XX
XX 20-APR-2000; 2000WO-CA000430.
XX
XX 23-APR-1999; 99US-0130508P.
XX (UYGU-) UNIV GUELPH.
XX Forsberg CW, Golovan S, Phillips JP;
XX WPI; 2000-687245/67.
XX N-PSDB; AAC68300.
XX Transgenic non-human animal for gastrointestinal tract specific
PT expression of a protein, preferably phytase, comprises a nucleic acid
PT sequence including a heterologous transgene construct encoding the
PT protein.
XX
XX Disclosure; Fig 23; 152pp; English.
XX
XX The present invention provides transgenic animals which produce desired
CC proteins, in this case pigs which expresses phytase in the salivary
CC gland. Low phytase production levels result in phytate in the diet being
CC excreted and causing phosphorus contamination in water, as well as
CC reducing the growth of animals. The invention provides a number of
CC transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC on 12-SEP-2003 to standardise OS field)
XX
SQ Sequence 432 AA;
Query Match 97.1%; Score 2235; DB 3; Length 432;
Best Local Similarity 99.5%; Pred. No. 1.6e-216;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MKAILPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDWP 60
Db 1 MKAILPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDWP 60
Qy 61 TWPVKLGLTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGOVAIIADVDERTRKTGE 120
Db 61 TWPVKLGLTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGOVAIIADVDERTRKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAIILSRAGGSIA DPTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAIILSRAGGSIA DPTGH 180
Qy 181 RQTAFLRELRLVLPFQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 RQTAFLRELRLVLPFQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLLQQAQMPGPGWGRITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Db 241 EIFLLQQAQMPGPGWGRITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Qy 301 LTPHPKQAYGVTLPTSVLFIAGHDTNLNLGGALELNWTLPGQPDNTPPGGELVPERW 360
Db 301 LTPHPKQAYGVTLPTSVLFIAGHDTNLNLGGALELNWTLPGQPDNTPPGGELVPERW 360
Qy 361 RRLSDNSQWIVQSVLFTLQMRDKTPLSLNTPPEVKLTLAGCEERNAQCMCSLAGFTQ 420
Db 361 RRLSDNSQWIVQSVLFTLQMRDKTPLSLNTPPEVKLTLAGCEERNAQCMCSLAGFTQ 420
Qy 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432

RESULT 15
AAB36260
ID AAB36260 standard; protein; 432 AA.
XX
AC AAB36260;
XX
DT 12-SEP-2003 (revised)
DT 20-FEB-2001 (first entry)
XX
DE R15/APPA plasmid translated sequence.
XX
KW Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KW environmental pollution; pig.
XX
OS Rattus sp.
OS Escherichia coli.
OS Chimeric.
XX
PN WO200064247-A1.
XX
PD 02-NOV-2000.
XX
PF 20-APR-2000; 2000WO-CA000430.
XX
PR 23-APR-1999; 99US-0130508P.
XX
PA (UYGU-) UNIV GUELPH.
XX
PI Foraberg CW, Golovan S, Phillips JP;
XX
XX WPI: 2000-687245/67.
DR N-PSDB; AAC68297.
XX

Transgenic non-human animal for gastrointestinal tract specific
expression of a protein, preferably phytase, comprises a nucleic acid
sequence including a heterologous transgene construct encoding the
protein.
XX
PS Disclosure; Fig 20; 152pp; English.
XX

The present invention provides transgenic animals which produce desired
proteins, in this case pigs which expresses phytase in the salivary
gland. Low phytase production levels result in phytate in the diet being
excreted and causing phosphorus contamination in water, as well as
reducing the growth of animals. The invention provides a number of
transgenes containing the E. coli APPA phytase coding sequence. (Updated
on 12-SEP-2003 to standardise OS field)
XX
SQ Sequence 432 AA;
Query Match 97.1%; Score 2235; DB 3; Length 432;
Best Local Similarity 99.5%; Pred. No. 1.6e-216;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MKAIIIPFLSLIPLTPOSAPAQSEPELKESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
DB 1 MKAIIIPFLSLIPLTPOSAPAQSEPELKESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
QY 61 TWPVKLGMLTPRGGLIAYLGHYQRLVADGLLAKKCPQSGQVAIIADVDERTKTGE 120
DB 61 TWPVKLGMLTPRGGLIAYLGHYQRLVADGLLAKKCPQSGQVAIIADVDERTKTGE 120
QY 121 APAAGLAPDCAITVHTQADTSSPDLFNLPLKTVGVOLDNANVTDAILSRAGGSIAFTGH 180
DB 121 APAAGLAPDCAITVHTQADTSSPDLFNLPLKTVGVOLDNANVTDAILSRAGGSIAFTGH 180
QY 181 RQTAFRELVNLPQSNICLKREKQDESCSLTQALPSELKVSADNVSITGAVSLASMLT 240
DB 181 RQTAFRELVNLPQSNICLKREKQDESCSLTQALPSELKVSADNVSITGAVSLASMLT 240
QY 241 EIFLLQQAQGMPEPCWGRIITDHSQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300

Search completed: June 13, 2006, 10:22:03
Job time : 141.266 secs

241 EIFLLQQAQGMPEPCWGRIITDHSQWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIKTA 300
301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
301 LTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGCEVKLTLAGCERNAQGMCSLAGFTQ 420
361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGCEVKLTLAGCERNAQGMCSLAGFTQ 420
421 IVNEARIPACSL 432
421 IVNEARIPACSL 432

THIS PAGE LEFT BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2006, 10:17:34 ; Search time 192.752 Seconds
(without alignments)
2111.554 Million cell updates/sec

Title: US-10-601-319-2

Perfect score: 2302

Sequence: 1 MKAILIPLSLILPLTQSA.....IVNEARIPACSLRSHHHHH 440

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_7.2.1*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2235	97.1	432	1	PPA_ECOLI
2	2235	97.1	432	2	Q3Z3E1 shigella so
3	2231	96.9	432	2	Q8GN88 ECOLI
4	2229	96.8	432	2	Q8KDK6 ECOLI
5	2229	96.8	432	2	Q8KDK7 ECOLI
6	2229	96.8	432	2	Q8KDK8 ECOLI
7	2229	96.8	432	2	Q8KDK9 ECOLI
8	2228	96.8	432	2	Q8KDK9 ECOLI
9	2228	96.8	432	2	Q8KDK9 ECOLI
10	2225	96.7	432	2	Q8KDK5 ECOLI
11	2222	96.5	432	2	Q8KDK2 shigella fl
12	2221	96.5	432	2	Q8KDK8 ECOLI
13	2218	96.4	432	2	Q7UD08 SHIFL
14	2213	96.1	446	2	Q8CW75 ECOL6
15	2210	96.0	434	2	Q7AFW3 ECO57
16	2210	96.0	444	2	Q8KDC29 ECO57
17	2191	95.2	434	2	Q3ZHS7 SHIDS
18	1374.5	59.7	433	2	Q676V7 CITFR
19	1361.5	59.1	433	2	Q2VY22 CITRAC
20	1077	45.8	444	2	Q6U677 9ENTR
21	1054.5	45.8	444	2	Q6TAQ8 9ENTR
22	943.5	41.0	441	2	Q669R3 YERPS
23	943.5	41.0	441	2	Q8ZFP6 YERPE
24	545	23.7	413	2	Q8XBZ6 ECO57
25	541	23.5	413	2	Q3Z3B9 SHISS
26	541	23.5	417	1	AGP PROR
27	539	23.4	413	1	AGP_ECOLI
28	538	23.4	413	1	Q3ZRV6 SHIFL
29	533	23.2	413	2	Q3ZHQ8 SHIDS
30	530	23.0	413	2	Q5PG89 SALPA
31	529	23.0	413	1	AGP_SALTY

32 529 23.0 413 2 Q57QN8 SALCH
33 526 22.8 414 2 Q9AAQ4 CAUCR
34 524 22.8 413 2 Q8Z7P1 SALT
35 523 22.7 413 2 Q6E1V9 ENTCL
36 505 21.9 392 2 Q8PP76 XANAC
37 505 21.9 443 2 Q4UR06 XANCS
38 505 21.9 443 2 Q8P330 XANCP
39 483 21.0 515 2 Q3BXB8 XANCS
40 480.5 20.9 435 2 Q8PF53 XANAC
41 476.5 20.7 433 2 Q3BM07 XANCS
42 474 20.6 532 2 Q5GW75 XANOR
43 467.5 20.3 435 2 Q6CZP4 ERWCT
44 439 19.1 427 2 Q4ZMU2 PSEU2
45 433.5 18.8 428 2 Q8GD20 PSESX

ALIGNMENTS

RESULT 1

ID_PPA_ECOLI STANDARD; PRT; 432 AA.
AC P07I02;
DT 01-APR-1988, integrated into UniProtKB/Swiss-Prot.
DT 01-AUG-1991, sequence version 2.
DT 07-MAR-2006, entry version 63.
DE Periplasmic appA protein precursor [Includes: Phosphoanhydride
DE phosphohydrolase (EC 3.1.3.2) (pH 2.5 acid phosphatase) (AP); 4-
DE Phytase (EC 3.1.3.26)].
GN Name=appA; OrderedLocNames=b0980;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND PROTEIN SEQUENCE OF 23-29.
RC STRAIN=K12;
RX MEDLINE=90368616; PubMed=2168385;
RA Dassa J., Marck C., Boquet P.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase.";
RL J. Bacteriol. 172:5497-5500(1990).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=K12 / W3110;
RX MEDLINE=97061202; PubMed=8905232; DOI=10.1093/dnares/3.3.137;
RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kaishimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nakamoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
RN [4]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 1-112.
RX MEDLINE=87271766; PubMed=3038201; DOI=10.1016/0300-9084(87)90045-9;
RA Touati E., Danchin A.;
RT "The structure of the promoter and amino terminal region of the pH 2.5
RT acid phosphatase structural gene (appA) of E. coli: a negative control
RT of transcription mediated by cyclic AMP.";

RL Biochimie 69:215-221(1987).
RN [5]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 1-17.
RC STRAIN=K12;
RA MEDLINE=92049231; PubMed=1658595; DOI=10.1007/BF00267454;
RA Dassa J., Fsihi H., Marck C., Dion M., Kieffer-Bontemps M.,
RA Boquet P.L.;
RT "A new oxygen-regulated operon in Escherichia coli comprises the genes
RT for a putative third cytochrome oxidase and for pH 2.5 acid
RT phosphatase (appa).";
RL Mol. Gen. Genet. 229:341-352(1991).
RN [6]
RP CHARACTERIZATION, AND PROTEIN SEQUENCE OF 23-34.
RA Greiner R., Jany K.-D.;
RT "Characterization of a phytase from Escherichia coli.";
RL Biol. Chem. Hoppe-Seyler 372:664-665(1991).
RN [7]
RP CHARACTERIZATION, AND PROTEIN SEQUENCE OF 23-35.
RX MEDLINE=93256556; PubMed=8387749; DOI=10.1006/abbi.1993.1261;
RA Greiner R., Konietzny U., Jany K.-D.;
RT "Purification and characterization of two phytases from Escherichia
RT coli.";
RL Arch. Biochem. Biophys. 303:107-113(1993).
RN [8]
RP CHARACTERIZATION.
RX MEDLINE=20161462; PubMed=10696472; DOI=10.1139/cjm-46-1-59;
RA Golovan S., Wang G., Zhang J., Forsberg C.W.;
RT "Characterization and overproduction of the Escherichia coli appA
RT encoded bifunctional enzyme that exhibits both phytase and acid
RT phosphatase activities.";
RL Can. J. Microbiol. 46:59-71(2000).
RN [9]
RP MUTAGENESIS.
RX MEDLINE=93054596; PubMed=1429631;
RA Ostalin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.-M.,
RA van Etten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase.";
RL J. Biol. Chem. 267:22830-22836(1992).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.05 ANGSTROMS).
RX MEDLINE=20122624; PubMed=10655611; DOI=10.1038/72371;
RA Lim D., Golovan S., Forsberg C.W., Jia Z.;
RT "Crystal structures of Escherichia coli phytase and its complex with
RT phytate.";
RL Nat. Struct. Biol. 7:108-113(2000).
CC -1- CATALYTIC ACTIVITY: A phosphate monoester + H(2)O = an alcohol +
CC phosphate.
CC -1- CATALYTIC ACTIVITY: Myo-inositol hexakisphosphate + H(2)O = 1D-
CC myo-inositol 1,2,3,4,5-pentakisphosphate + phosphate.
CC -1- SUBUNIT: Monomer.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- INDUCTION: In addition to CAMP-mediated control, this enzyme is
CC induced when bacterial cultures reach stationary phase; its
CC synthesis is triggered by phosphate starvation or a shift from
CC aerobic to anaerobic conditions.
CC -1- SIMILARITY: Belongs to the histidine acid phosphatase family.
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; M58708; AAA72086.1; -; Genomic_DNA.
CC EMBL; U00096; AAC74065.1; -; Genomic_DNA.
CC EMBL; D90735; BAA35745.1; -; Genomic_DNA.
CC EMBL; X05471; CAA29031.1; -; Genomic_DNA.
CC EMBL; M58708; -; NOT ANNOTATED CDS; Genomic_DNA.
CC EMBL; S63811; AAB20286.1; -; Genomic_DNA.
CC PIR; B36733; B36733.
CC PDB; 1DKU; X-ray; A/B=23-432.
CC PDB; 1DKM; X-ray; A=23-432.
CC PDB; 1DKN; X-ray; A=23-432.
CC PDB; 1DKO; X-ray; A=23-432.
CC PDB; 1DKP; X-ray; A=23-432.

DR PDB; 1DKO; X-ray; A=23-432.
DR SWISS-2DPAGE; P07102; COLI.
DR GenomeReviews; U00096_GR; b0980.
DR EchoBASE; EB0047; -.
DR EcoGene; EG10049; appA.
DR BioCyc; BcoCyc:APPA-MONOMER; -.
DR InterPro; IPR000560; Hisac_phsphtse.
DR Pfam; PF003328; Acid_phosphat_A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW 3D-structure; Complete proteome; Direct protein sequencing; Hydrolase;
KW Multifunctional enzyme; Periplasmic; Signal.
FT SIGNAL 1 22
FT CHAIN 23 432 Periplasmic appA protein.
FT /FTID=PRO_0000023947.
FT Nucleophile.
FT Proton donor.
FT ACT_SITE 39 39
FT ACT_SITE 326 326
FT DISULFID 99 130
FT DISULFID 155 430
FT DISULFID 200 210
FT DISULFID 404 413
FT CONFLICT 51 66 MODVTPDAWPTWPVKL -> NAGCHPERMANLAGKT (in
FT Ref. 3).
FT CONFLICT 75 76 EL -> DV (in Ref. 4).
FT CONFLICT 112 112 D -> S (in Ref. 4).
FT STRAND 27 38
FT STRAND 41 42
FT STRAND 45 45
FT HELIX 49 53
FT TURN 54 54
FT STRAND 56 57
FT STRAND 63 64
FT TURN 66 67
FT STRAND 69 69
FT STRAND 71 90
FT TURN 91 92
FT STRAND 93 94
FT STRAND 96 98
FT TURN 102 104
FT STRAND 105 109
FT STRAND 111 112
FT HELIX 113 126
FT STRAND 127 127
FT TURN 128 129
FT STRAND 134 135
FT STRAND 138 138
FT TURN 140 141
FT STRAND 142 142
FT HELIX 145 147
FT TURN 149 153
FT STRAND 154 155
FT HELIX 159 169
FT TURN 170 172
FT HELIX 174 179
FT TURN 180 181
FT HELIX 182 192
FT TURN 193 193
FT HELIX 194 196
FT STRAND 197 197
FT HELIX 198 201
FT TURN 202 202
FT STRAND 204 204
FT TURN 206 207
FT STRAND 208 208
FT HELIX 212 215
FT STRAND 216 216
FT STRAND 220 222
FT STRAND 224 225
FT STRAND 227 229
FT TURN 231 231
FT HELIX 232 249
FT TURN 250 250
FT STRAND 252 252

FT HELIX 254 257
FT TURN 258 258

Query Match
Best Local Similarity 97.1%; Score 2235; DB 1; Length 432;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKAIIIPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
DB 1 MKAIIIPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPDAMP 60

QY 61 TWPVKLGLTTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
DB 61 TWPVKLGLTTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGVAIADVDERTKTGE 120

QY 121 APAAGLAPDCAITVHTQADTSSPDLFNLKGTGVCQDNLNANTDAILSRAGGSIAADFTGH 180
DB 121 APAAGLAPDCAITVHTQADTSSPDLFNLKGTGVCQDNLNANTDAILSRAGGSIAADFTGH 180

QY 181 ROTAFRELERVLNFPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFRELERVLNFPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

QY 241 EIFLLQQAQGMPEPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLDLIKTA 300
DB 241 EIFLLQQAQGMPEPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLDLIKTA 300

QY 301 LTPHPQQAQYGVTLPTSFLVFIAGHDTNLNLAGALELNTLPGQPDNTPPGGLVFERW 360
DB 301 LTPHPQQAQYGVTLPTSFLVFIAGHDTNLNLAGALELNTLPGQPDNTPPGGLVFERW 360

QY 361 RRLSDNSQWIVSLVFQTLQMRDKTPLSLNTPPGEVKLTLAGCBERNAQMCSLAGFTQ 420
DB 361 RRLSDNSQWIVSLVFQTLQMRDKTPLSLNTPPGEVKLTLAGCBERNAQMCSLAGFTQ 420

QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 2

Q323E1_SHISS PRELIMINARY; PRT; 432 AA.
AC Q323E1;
DT 27-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 27-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 5.
DE Phosphoanhydride phosphorylase.
GN Name=appa; OrderedLocustNames=SSO_0987; ORFNames=SSO_0987;
OS Shigella sonnei (strain S8046).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=300269;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=1625786; DOI=10.1093/nar/gki954;
RA Yang F., Yang J., Zhang X., Chen L., Jiang Y., Yan Y., Tang X.,
RA Wang J., Xiong Z., Dong J., Xu Y., Xu X., Sun L., Chen S.,
RA Nie H., Peng J., Xu J., Wang Y., Yuan Z., Wen Y., Yao Z., Shen Y.,
RA Qiang B., Hou Y., Yu J., Jin Q.,
RT "Genome dynamics and diversity of Shigella species, the etiologic
RT agents of bacillary dysentery."
RL Nucleic Acids Res. 33:6445-6458(2005).
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
CC EMBL; CP000338; AA287721.1; -; Genomic DNA.
CC GO; GO:0003993; P:acid phosphatase activity; IEA.
CC InterPro; IPR000560; HisAc.phaphtase.
CC Pfam; PF00328; Acid phosphatase A; 1.
CC PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
CC PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.

KW Complete proteome.
SQ SEQUENCE 432 AA; 47057 MW; 6510C6C579177F11 CRC64;

Query Match
Best Local Similarity 97.1%; Score 2235; DB 2; Length 432;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKAIIIPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
DB 1 MKAIIIPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPDAMP 60

QY 61 TWPVKLGLTTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
DB 61 TWPVKLGLTTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGVAIADVDERTKTGE 120

QY 121 APAAGLAPDCAITVHTQADTSSPDLFNLKGTGVCQDNLNANTDAILSRAGGSIAADFTGH 180
DB 121 APAAGLAPDCAITVHTQADTSSPDLFNLKGTGVCQDNLNANTDAILSRAGGSIAADFTGH 180

QY 181 ROTAFRELERVLNFPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFRELERVLNFPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

QY 241 EIFLLQQAQGMPEPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLDLIKTA 300
DB 241 EIFLLQQAQGMPEPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLDLIKTA 300

QY 301 LTPHPQQAQYGVTLPTSFLVFIAGHDTNLNLAGALELNTLPGQPDNTPPGGLVFERW 360
DB 301 LTPHPQQAQYGVTLPTSFLVFIAGHDTNLNLAGALELNTLPGQPDNTPPGGLVFERW 360

QY 361 RRLSDNSQWIVSLVFQTLQMRDKTPLSLNTPPGEVKLTLAGCBERNAQMCSLAGFTQ 420
DB 361 RRLSDNSQWIVSLVFQTLQMRDKTPLSLNTPPGEVKLTLAGCBERNAQMCSLAGFTQ 420

QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 3
Q8GN88_ECOLI PRELIMINARY; PRT; 432 AA.
AC Q8GN88;
DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2003, sequence version 1.
DT 07-FEB-2006, entry version 11.
DE AppA.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Zhang Z., He J., Yao B., Zhou Y., Chen Y., Yi Y.,
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Chen Y., Zhu Z., Zhang Z., He J.,
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
CC EMBL; AF537219; AAN28334.1; -; Genomic DNA.
CC HSSP; P07102; 1DKM.
CC SMR; Q8GN88; 23-432.
CC GO; GO:0003993; P:acid phosphatase activity; IEA.
CC InterPro; IPR000560; HisAc.phaphtase.
CC Pfam; PF00328; Acid phosphatase A; 1.
CC PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
CC PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.

```
SQ SEQUENCE 432 AA; 47056 MW; 5B355D7687773737 CRC64;

Query Match 96.9%; Score 2231; DB 2; Length 432;
Best Local Similarity 99.3%; Pred. No. 1.5e-162;
Matches 429; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
DB 1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60

QY 61 TWPVKGLWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGOVAILIADVDERTRKTGE 120
DB 61 TWPVKGLWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGOVAILIADVDERTRKTGE 120

QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQLDNANVTDAILSRAGGSIADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQLDNANVTDAILSRAGGSIADFTGH 180

QY 181 ROTAFRELERVLNFPQSNCLCKREKQDESCSLTOALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFRELERVLNFPQSNCLCKREKQDESCSLTOALPSELKVSADNVSLTGAVSLASMLT 240

QY 241 EIFLLOAQOMPEPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLDLLIMAA 300
DB 241 EIFLLOAQOMPEPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLDLLIMAA 300

QY 301 LTPHPKQKQAYGVTLPTSVLFIAGHDTNLNLGGALELNWTLPGQPDNTPPGSELVPERW 360
DB 301 LTPHPKQKQAYGVTLPTSVLFIAGHDTNLNLGGALELNWTLPGQPDNTPPGSELVPERW 360

QY 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGVKLTLAGCEERNAQMGCSLAGFTQ 420
DB 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGVKLTLAGCEERNAQMGCSLAGFTQ 420

QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 4
Q8RKD6_ECOLI
ID Q8RKD6_ECOLI PRELIMINARY; PRT; 432 AA.
AC Q8RKD6;
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2002, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Periplasmic phosphoanhydride phosphohydrolase.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=90368616; PubMed=2168385;
RA Dassa J., Marck C., Boquet P.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase."
RL J. Bacteriol. 172:5497-5500(1990).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93054596; PubMed=1429631;
RA Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.-M.,
RA van Eeten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase."
RL J. Biol. Chem. 267:22830-22836(1992).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DB EMBL; L03374; AAA00006.1; -; Genomic_DNA.
```

```
DR HSSP: P07102; 1DKM.
DR SMR; Q8RKD6; 23-432.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000560; HisAc phsphtse.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolase.
SQ SEQUENCE 432 AA; 46972 MW; AF86C41EA6193AC5 CRC64;

Query Match 96.8%; Score 2229; DB 2; Length 432;
Best Local Similarity 99.3%; Pred. No. 2.2e-162;
Matches 429; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
DB 1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60

QY 61 TWPVKGLWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGOVAILIADVDERTRKTGE 120
DB 61 TWPVKGLWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGOVAILIADVDERTRKTGE 120

QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQLDNANVTDAILSRAGGSIADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQLDNANVTDAILSRAGGSIADFTGH 180

QY 181 ROTAFRELERVLNFPQSNCLCKREKQDESCSLTOALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFRELERVLNFPQSNCLCKREKQDESCSLTOALPSELKVSADNVSLTGAVSLASMLT 240

QY 241 EIFLLOAQOMPEPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLDLLIMAA 300
DB 241 EIFLLOAQOMPEPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLDLLIMAA 300

QY 301 LTPHPKQKQAYGVTLPTSVLFIAGHDTNLNLGGALELNWTLPGQPDNTPPGSELVPERW 360
DB 301 LTPHPKQKQAYGVTLPTSVLFIAGHDTNLNLGGALELNWTLPGQPDNTPPGSELVPERW 360

QY 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGVKLTLAGCEERNAQMGCSLAGFTQ 420
DB 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGVKLTLAGCEERNAQMGCSLAGFTQ 420

QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 5
Q8RKD7_ECOLI
ID Q8RKD7_ECOLI PRELIMINARY; PRT; 432 AA.
AC Q8RKD7;
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2002, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Periplasmic phosphoanhydride phosphohydrolase.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=90368616; PubMed=2168385;
RA Dassa J., Marck C., Boquet P.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase."
RL J. Bacteriol. 172:5497-5500(1990).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93054596; PubMed=1429631;
RA Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.-M.,
```

RA van Etten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase."
RL J. Biol. Chem. 267:22830-22836(1992).
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC
CC EMBL: L03373; AAA00005.1; -; Genomic_DNA.
DR HSSP: P07102; 1DKM.
DR SMR: Q8RKH7; 23-432.
DR GO: GO:0003993; F:acid phosphatase activity; IEA.
DR GO: GO:0016787; F:hydrolyase activity; IEA.
DR InterPro: IPR000560; H:ac phosphatase.
DR Pfam: PF00328; Acid phosphat A; 1.
DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolyase.
SQ SEQUENCE 432 AA; 46972 MW; 755D5E4B1AD916A6 CRC64;

Query Match 96.8%; Score 2229; DB 2; Length 432;
Best Local Similarity 99.3%; Pred. No. 2.2e-162;
Matches 429; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKAILIPFLSLIPLTPQSAFAQSEBELKESVIVSRHGVAAPTKATQLMQDVTDPDWP 60
DB 1 MKAILIPFLSLIPLTPQSAFAQSEBELKESVIVSRHGVAAPTKATQLMQDVTDPDWP 60
QY 61 TWPVGLGWLTPRGGLIAYLGHYQORLVADGLAKKCGPQSGOVAIADVERTRTKGE 120
DB 61 TWPVGLGWLTPRGGLIAYLGHYQORLVADGLAKKCGPQSGOVAIADVERTRTKGE 120
QY 121 AFAAGLADCAITVHTQADTSSPDPLFNPPLKTVGVCQLDNANVTDAIISRAGSIADFTGH 180
DB 121 AFAAGLADCAITVHTQADTSSPDPLFNPPLKTVGVCQLDNANVTDAIISRAGSIADFTGH 180
QY 181 RQTAFLRELRLVNFQSNCLCKREKODESCSLTQALPSELKVSADNVSLTGAVSIASMLT 240
DB 181 RQTAFLRELRLVNFQSNCLCKREKODESCSLTQALPSELKVSADNVSLTGAVSIASMLT 240
QY 241 EIFLLQQAQMPBEQWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSATPLDLIKYA 300
DB 241 EIFLLQQAQMPBEQWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSATPLDLIKYA 300
QY 301 LTPHPQQAQVAVTLPTSVLFIAGHDYMLANLGGALBELNMTLPQGPDMTPPGGELVFERW 360
DB 301 LTPHPQQAQVAVTLPTSVLFIAGHDYMLANLGGALBELNMTLPQGPDMTPPGGELVFERW 360
QY 361 RRLSDNSQMIQVSLVFQTLQQRDKTPLSLNTPPGEVKTLLAGCBERRNAQMCISLAGFTQ 420
DB 361 RRLSDNSQMIQVSLVFQTLQQRDKTPLSLNTPPGEVKTLLAGCBERRNAQMCISLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 6
Q8RKH8_ECOLI PRELIMINARY; PRT; 432 AA.
AC Q8RKH8;
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 13.
DE Periplasmic phosphonhydride phosphohydrolase.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=9036616; PubMed=2168385;
RX

RA Dasas J., Marck C., Boquet P.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appa
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase."
RL J. Bacteriol. 172:5497-5500(1990).
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC
CC EMBL: L03372; AAA00004.1; -; Genomic_DNA.
DR HSSP: P07102; 1DKM.
DR SMR: Q8RKH8; 23-432.
DR GO: GO:0003993; F:acid phosphatase activity; IEA.
DR GO: GO:0016787; F:hydrolyase activity; IEA.
DR InterPro: IPR000560; H:ac phosphatase.
DR Pfam: PF00328; Acid phosphat A; 1.
DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolyase.
SQ SEQUENCE 432 AA; 46972 MW; 9A85536B57FCCFB5 CRC64;

Query Match 96.8%; Score 2229; DB 2; Length 432;
Best Local Similarity 99.3%; Pred. No. 2.2e-162;
Matches 429; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKAILIPFLSLIPLTPQSAFAQSEBELKESVIVSRHGVAAPTKATQLMQDVTDPDWP 60
DB 1 MKAILIPFLSLIPLTPQSAFAQSEBELKESVIVSRHGVAAPTKATQLMQDVTDPDWP 60
QY 61 TWPVGLGWLTPRGGLIAYLGHYQORLVADGLAKKCGPQSGOVAIADVERTRTKGE 120
DB 61 TWPVGLGWLTPRGGLIAYLGHYQORLVADGLAKKCGPQSGOVAIADVERTRTKGE 120
QY 121 AFAAGLADCAITVHTQADTSSPDPLFNPPLKTVGVCQLDNANVTDAIISRAGSIADFTGH 180
DB 121 AFAAGLADCAITVHTQADTSSPDPLFNPPLKTVGVCQLDNANVTDAIISRAGSIADFTGH 180
QY 181 RQTAFLRELRLVNFQSNCLCKREKODESCSLTQALPSELKVSADNVSLTGAVSIASMLT 240
DB 181 RQTAFLRELRLVNFQSNCLCKREKODESCSLTQALPSELKVSADNVSLTGAVSIASMLT 240
QY 241 EIFLLQQAQMPBEQWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSATPLDLIKYA 300
DB 241 EIFLLQQAQMPBEQWGRITDSHQWNTLLSLHNAQFYLLQRTPEVARSATPLDLIKYA 300
QY 301 LTPHPQQAQVAVTLPTSVLFIAGHDYMLANLGGALBELNMTLPQGPDMTPPGGELVFERW 360
DB 301 LTPHPQQAQVAVTLPTSVLFIAGHDYMLANLGGALBELNMTLPQGPDMTPPGGELVFERW 360
QY 361 RRLSDNSQMIQVSLVFQTLQQRDKTPLSLNTPPGEVKTLLAGCBERRNAQMCISLAGFTQ 420
DB 361 RRLSDNSQMIQVSLVFQTLQQRDKTPLSLNTPPGEVKTLLAGCBERRNAQMCISLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 7
Q8RKH0_ECOLI PRELIMINARY; PRT; 432 AA.
AC Q8RKH0;
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 13.
DE Periplasmic phosphonhydride phosphohydrolase.

GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=9036816; PubMed=2168385;
RA Dasra J., Marck C., Boquet P.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase.";
RL J. Bacteriol. 172:5497-5500 (1990).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93054596; PubMed=1429631;
RA Ostaman K., Harms E.H., Stevies P.B., Kuciel R., Zhou M.-M.,
RA van Echten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase.";
RL J. Biol. Chem. 267:22830-22836 (1992).
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; L03370; AAA00002.1; -; Genomic_DNA.
DR HSSP; P07102; 1DKM.
DR SMR; Q8RKE0; 23-432.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000560; HisGac_phosphatase.
DR Pfam; PF00328; Acid_phosphatase_A; 1.
DR PROSITE; PS00776; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolase.
SQ SEQUENCE 432 AA; 46972 MW; 5BBD632D4682EABF CRC64;
Query Match 96.8%; Score 2229; DB 2; Length 432;
Best Local Similarity 99.3%; Pred. No. 2.2e-162;
Matches 429; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MKAILIPFLSLILPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQTMQDVTTPDAMP 60
DB 1 MKAILIPFLSLILPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQTMQDVTTPDAMP 60
QY 61 TWPKLGMLTPRGGEIAYIGHYQORQLVADGLAKKGCPSQGVAILIADVDETRKTGE 120
DB 61 TWPKLGMLTPRGGEIAYIGHYQORQLVADGLAKKGCPSQGVAILIADVDETRKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOLDNANVTDAISRAGSIADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOLDNANVTDAISRAGSIADFTGH 180
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOLDNANVTDAISRAGSIADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOLDNANVTDAISRAGSIADFTGH 180
QY 181 RQTAFRELERVLNFPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 RQTAFRELERVLNFPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLQQAQGMPEPGWRITDSHOWNTLSLHNAQFYLLQRTPEVARSRATPLIDLIMAA 300
DB 241 EIFLLQQAQGMPEPGWRITDSHOWNTLSLHNAQFYLLQRTPEVARSRATPLIDLIMAA 300
QY 301 LTPHPQKQAGVTLPTSVLFIAGHDNTLANIGALBLNMTLPQDPNTPPGGELVPERW 360
DB 301 LTPHPQKQAGVTLPTSVLFIAGHDNTLANIGALBLNMTLPQDPNTPPGGELVPERW 360
QY 361 RLSDNSQMWIQVSLVFTLQGMKDKTPLSLNTPGGEVYKTLTLAGCEERNAQMGCSLAFQTQ 420
DB 361 RLSDNSQMWIQVSLVFTLQGMKDKTPLSLNTPGGEVYKTLTLAGCEERNAQMGCSLAFQTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 8
Q31YP3 SHIBS
ID Q31YP3_SHIBS PRELIMINARY; PRT; 432 AA.
AC Q31YP3;
DT 06-DEC-2005, integrated into UniProtKB/TrEMBL.
DT 06-DEC-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Phosphoanhydride phosphorylase.
GN Name=appa; OrderedLocustNames=SBO_2250;
OS Shigella boydii serotype 4 (strain Sb227).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=300268;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=16275786; DOI=10.1093/nar/gki1954;
RA Yang F., Yang J., Zhang X., Chen L., Ujang Y., Yan Y., Tang X.,
RA Wang J., Xiong Z., Dong J., Xue Y., Zhu Y., Xu X., Sun L., Chen S.,
RA Nie H., Peng J., Xu J., Wang Y., Yuan Z., Wen Y., Yao Z., Shen Y.,
RA Qiang B., Hou Y., Yu J., Jin Q.;
RT "Genome dynamics and diversity of Shigella species, the etiologic
RT agents of bacillary dysentery.";
RL Nucleic Acids Res. 33:6445-6458 (2005).
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; CP000036; ABB66815.1; -; Genomic_DNA.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR Complete proteome.
SQ SEQUENCE 432 AA; 47064 MW; 8569022BA28D9C73 CRC64;
Query Match 96.8%; Score 2228; DB 2; Length 432;
Best Local Similarity 99.3%; Pred. No. 2.6e-162;
Matches 429; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MKAILIPFLSLILPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQTMQDVTTPDAMP 60
DB 1 MKAILIPFLSLILPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQTMQDVTTPDAMP 60
QY 61 TWPKLGMLTPRGGEIAYIGHYQORQLVADGLAKKGCPSQGVAILIADVDETRKTGE 120
DB 61 TWPKLGMLTPRGGEIAYIGHYQORQLVADGLAKKGCPSQGVAILIADVDETRKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOLDNANVTDAISRAGSIADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOLDNANVTDAISRAGSIADFTGH 180
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOLDNANVTDAISRAGSIADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOLDNANVTDAISRAGSIADFTGH 180
QY 181 RQTAFRELERVLNFPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 RQTAFRELERVLNFPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLQQAQGMPEPGWRITDSHOWNTLSLHNAQFYLLQRTPEVARSRATPLIDLIMAA 300
DB 241 EIFLLQQAQGMPEPGWRITDSHOWNTLSLHNAQFYLLQRTPEVARSRATPLIDLIMAA 300
QY 301 LTPHPQKQAGVTLPTSVLFIAGHDNTLANIGALBLNMTLPQDPNTPPGGELVPERW 360
DB 301 LTPHPQKQAGVTLPTSVLFIAGHDNTLANIGALBLNMTLPQDPNTPPGGELVPERW 360
QY 361 RLSDNSQMWIQVSLVFTLQGMKDKTPLSLNTPGGEVYKTLTLAGCEERNAQMGCSLAFQTQ 420
DB 361 RLSDNSQMWIQVSLVFTLQGMKDKTPLSLNTPGGEVYKTLTLAGCEERNAQMGCSLAFQTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 9
Q8RKD9_ECOLI
ID Q8RKD9_ECOLI PRELIMINARY; PRT; 432 AA.
AC Q8RKD9;
Q8RKD9;

DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2002, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Periplasmic phosphoanhydride phosphohydrolase.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=9036816; PubMed=2168385;
RA Dassa J., Marc C., Boquet P.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase.";
RL J. Bacteriol. 172:5497-5500(1990).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93054596; PubMed=1429631;
RA Ostalan K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.-M.,
RA van Etten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase.";
RL J. Biol. Chem. 267:22830-22836(1992).
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs license
DR EMBL, L03371; AAA00003.1; -; Genomic_DNA.
DR HSSP; P07102; 1DKO.
DR SMR; O8RKD9; 23-432.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000560; HsAc_phsphtse.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolyase.
SQ SEQUENCE 432 AA; 47034 MW; 9F29B9DF9C368175 CRC64;

Query Match 96.8%; Score 2228; DB 2; Length 432;
Best Local Similarity 99.3%; Pred. No. 2,6e-162;
Matches 429; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKAILIPFLSLILPLTPQSAFAQSEPELKESVIVSRHGVRATPKATQLMQDVTTPDMP 60
DB 1 MKAILIPFLSLILPLTPQSAFAQSEPELKESVIVSRHGVRATPKATQLMQDVTTPDMP 60

QY 61 TWPVKLGMLTPRGSELIAVYGHYGRORLVADGLAKKGCPSGGVAIITADVDERTRKTGE 120
DB 61 TWPVKLGMLTPRGSELIAVYGHYGRORLVADGLAKKGCPSGGVAIITADVDERTRKTGE 120

QY 121 AFAAGLAPDCAITVHTQADTSSPPLFNPPLKTVGCQLDNANVTDAIISRAGGSIAADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPPLFNPPLKTVGCQLDNANVTDAIISRAGGSIAADFTGH 180

QY 121 AFAAGLAPDCAITVHTQADTSSPPLFNPPLKTVGCQLDNANVTDAIISRAGGSIAADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPPLFNPPLKTVGCQLDNANVTDAIISRAGGSIAADFTGH 180

QY 181 ROTAFRELERLVANPQSNCLKREKODSCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFRELERLVANPQSNCLKREKODSCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

QY 241 EIFLLQQAQGMPEBGMGRITDISHQNTLLSLHNAQFYLLQRTPEVARSRAATPLDLITMAA 300
DB 241 EIFLLQQAQGMPEBGMGRITDISHQNTLLSLHNAQFYLLQRTPEVARSRAATPLDLITMAA 300

QY 301 LTPHPQKQAVGVTLPTSVLFIAGHDTNLANLGALIELNMTLPQGPNDTTPPGSELVFERW 360
DB 301 LTPHPQKQAVGVTLPTSVLFIAGHDTNLANLGALIELNMTLPQGPNDTTPPGSELVFERW 360

QY 361 RLISNSNGWQVSLVFTQLQQRDKTPLYSLNTPPGSEVTLTLAGCERNAAQMGCSIAFTQ 420
DB 361 RLISNSNGWQVSLVFTQLQQRDKTPLYSLNTPPGSEVTLTLAGCERNAAQMGCSIAFTQ 420

QY 421 IVNEARIPACSL 432

DB 421 IVNEARIPACSL 432

RESULT 10
O8RKD5_ECOLI PRELIMINARY; PRT; 432 AA.
AC O8RKD5_5
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2002, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Periplasmic phosphoanhydride phosphohydrolase.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=9036816; PubMed=2168385;
RA Dassa J., Marc C., Boquet P.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase.";
RL J. Bacteriol. 172:5497-5500(1990).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93054596; PubMed=1429631;
RA Ostalan K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.-M.,
RA van Etten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase.";
RL J. Biol. Chem. 267:22830-22836(1992).
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs license
DR EMBL, L03375; AAA00007.1; -; Genomic_DNA.
DR HSSP; P07102; 1DKO.
DR SMR; O8RKD5; 23-432.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000560; HsAc_phsphtse.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
KW Hydrolyase.
SQ SEQUENCE 432 AA; 46991 MW; 951F393BA9A1A47C CRC64;

Query Match 96.7%; Score 2225; DB 2; Length 432;
Best Local Similarity 99.3%; Pred. No. 4,5e-162;
Matches 429; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKAILIPFLSLILPLTPQSAFAQSEPELKESVIVSRHGVRATPKATQLMQDVTTPDMP 60
DB 1 MKAILIPFLSLILPLTPQSAFAQSEPELKESVIVSRHGVRATPKATQLMQDVTTPDMP 60

QY 61 TWPVKLGMLTPRGSELIAVYGHYGRORLVADGLAKKGCPSGGVAIITADVDERTRKTGE 120
DB 61 TWPVKLGMLTPRGSELIAVYGHYGRORLVADGLAKKGCPSGGVAIITADVDERTRKTGE 120

QY 121 AFAAGLAPDCAITVHTQADTSSPPLFNPPLKTVGCQLDNANVTDAIISRAGGSIAADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPPLFNPPLKTVGCQLDNANVTDAIISRAGGSIAADFTGH 180

QY 121 AFAAGLAPDCAITVHTQADTSSPPLFNPPLKTVGCQLDNANVTDAIISRAGGSIAADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPPLFNPPLKTVGCQLDNANVTDAIISRAGGSIAADFTGH 180

QY 181 ROTAFRELERLVANPQSNCLKREKODSCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFRELERLVANPQSNCLKREKODSCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

QY 241 EIFLLQQAQGMPEBGMGRITDISHQNTLLSLHNAQFYLLQRTPEVARSRAATPLDLITMAA 300
DB 241 EIFLLQQAQGMPEBGMGRITDISHQNTLLSLHNAQFYLLQRTPEVARSRAATPLDLITMAA 300

QY 301 LTPHPQKQAVGVTLPTSVLFIAGHDTNLANLGALIELNMTLPQGPNDTTPPGSELVFERW 360

```

Db      301 LTPHPQKQAYGVTLPSPSVLFIAGADNTNLNLGALIELNMTLPGQPDNTPRGSELVFERW 360
Qy      361 RRLSDNSQMWIQVSLVFQTLQOMRDKTPLSLNTPRGSEVKTLLAGCEERNAQMGSLAGFTQ 420
Db      361 RRLSDNSQMWIQVSLVFQTLQOMRDKTPLSLNTPRGSEVKTLLAGCEERNAQMGSLAGFTQ 420
Qy      421 IYNEARIIPACSL 432
Db      421 IYNEARIIPACSL 432

RESULT 11
083RW2 SHIFL
ID 083RW2 SHIFL PRELIMINARY; PRT; 432 AA.
AC 083RW2;
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2005, sequence version 4.
DT 07-FEB-2006, entry version 15.
DE Phosphoanhydride phosphorylase; PH 2.5 acid phosphatase.
GN Name=appa; OrderedLocustNames=SF0982;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OC NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA).
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
CC EMBL; AE005674; AAN42610.2; -; Genomic_DNA.
CC SMR; Q83RW2; 23-432.
CC DR BIOCYC; SFLB198214; AAN42610.1-MONOMER; -;
CC DR GO; GO:0003993; F:acid phosphatase activity; IEA.
CC DR InterPro; IPR000560; HisAc_phosphatase.
CC DR Pfam; PF00328; Acid_phosphat_A; 2.
CC DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN 1.
CC DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN 1.
CC KW Complete proteome.
SQ SEQUENCE 432 AA; 47106 MW; 25093A5392B9C18C CRC64;

Query Match 96.5%; Score 2222; DB 2; Length 432;
Best Local Similarity 98.8%; Pred. No. 7.6e-162;
Matches 427; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy      1 MKAILIPFSLILPLTPQSAFAQSEPELKLESVIVYSRHGVRAPTKATQLMQDVTTPAMP 60
Db      1 MKAILIPFSLILPLTPQSAFAQSEPELKLESVIVYSRHGVRAPTKATQLMQDVTTPAMP 60
Qy      61 TWPKVGLMTLPRGSELIAVLGHYQORLVADGLAKKGCPSQGVALLADVDETRTKTGE 120
Db      61 TWPKVGLMTLPRGSELIAVLGHYQORLVADGLAKKGCPSQGVALLADVDETRTKTGE 120
Qy      121 TWPKVGLMTLPRGSELIAVLGHYQORLVADGLAKKGCPSQGVALLADVDETRTKTGE 180
Db      121 TWPKVGLMTLPRGSELIAVLGHYQORLVADGLAKKGCPSQGVALLADVDETRTKTGE 180
Qy      121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQLDNNAVTDAILSRAGGSIAFTGH 180
Db      121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQLDNNAVTDAILSRAGGSIAFTGH 180
Qy      181 RQTAFRELERVLNPFQSNCLNREKODESCSLTQALPSELKVADNVSLTGAVSLASMLT 240
Db      181 RQTAFRELERVLNPFQSNCLNREKODESCSLTQALPSELKVADNVSLTGAVSLASMLT 240
Qy      241 EIPLLQQAQGMPEBQWGRITDSHQMNTLLSLHNAQFYLLQRTPEVARSRATPLDLIMAA 300

```

```

Db      241 EIPLLQQAQGMPEBQWGRITDSHQMNTLLSLHNAQFYLLQRTPEVARSRATPLDLIMAA 300
Qy      301 LTPHPQKQAYGVTLPSPSVLFIAGDNTNLNLGALIELNMTLPGQPDNTPRGSELVFERW 360
Db      301 LTPHPQKQAYGVTLPSPSVLFIAGDNTNLNLGALIELNMTLPGQPDNTPRGSELVFERW 360
Qy      361 RRLSDNSQMWIQVSLVFQTLQOMRDKTPLSLNTPRGSEVKTLLAGCEERNAQMGSLAGFTQ 420
Db      361 RRLSDNSQMWIQVSLVFQTLQOMRDKTPLSLNTPRGSEVKTLLAGCEERNAQMGSLAGFTQ 420
Qy      421 IYNEARIIPACSL 432
Db      421 IYNEARIIPACSL 432

RESULT 12
06RK08 ECOLI
ID 06RK08 ECOLI PRELIMINARY; PRT; 432 AA.
AC 06RK08;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE Acid phosphatase/phytase 2.
GN Name=appa2;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OC NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99194564; PubMed=10092520; DOI=10.1006/dbrc.1999.0361;
RA Rodriguez B., Han Y., Lei X.G.;
RT "Cloning, sequencing, and expression of an Escherichia coli acid
RT phosphatase/phytase gene (appa2) isolated from pig colon."
RL Biochem. Biophys. Res. Commun. 257:117-123(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Lei X.G.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
CC EMBL; AY46073; AAR87658.1; -; Genomic_DNA.
CC SMR; 06RK08; 23-432.
CC DR GO; GO:0003993; F:acid phosphatase activity; IEA.
CC DR InterPro; IPR000560; HisAc_phosphatase.
CC DR Pfam; PF00328; Acid_phosphat_A; 1.
CC DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
CC DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN 1.
CC FT CHAIN 20 432 acid phosphatase/phytase 2.
SQ SEQUENCE 432 AA; 47042 MW; 71B05EBD2EA2674 CRC64;

Query Match 96.5%; Score 2221; DB 2; Length 432;
Best Local Similarity 99.1%; Pred. No. 9.1e-162;
Matches 428; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      1 MKAILIPFSLILPLTPQSAFAQSEPELKLESVIVYSRHGVRAPTKATQLMQDVTTPAMP 60
Db      1 MKAILIPFSLILPLTPQSAFAQSEPELKLESVIVYSRHGVRAPTKATQLMQDVTTPAMP 60
Qy      61 TWPKVGLMTLPRGSELIAVLGHYQORLVADGLAKKGCPSQGVALLADVDETRTKTGE 120
Db      61 TWPKVGLMTLPRGSELIAVLGHYQORLVADGLAKKGCPSQGVALLADVDETRTKTGE 120
Qy      121 TWPKVGLMTLPRGSELIAVLGHYQORLVADGLAKKGCPSQGVALLADVDETRTKTGE 180
Db      121 TWPKVGLMTLPRGSELIAVLGHYQORLVADGLAKKGCPSQGVALLADVDETRTKTGE 180
Qy      121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQLDNNAVTDAILSRAGGSIAFTGH 180
Db      121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQLDNNAVTDAILSRAGGSIAFTGH 180
Qy      181 RQTAFRELERVLNPFQSNCLNREKODESCSLTQALPSELKVADNVSLTGAVSLASMLT 240
Db      181 RQTAFRELERVLNPFQSNCLNREKODESCSLTQALPSELKVADNVSLTGAVSLASMLT 240

```

```
QY 241 EIFFLQQAQMPBPBGMRITDTSKQNTLLSLHNAQFYLLQRTPEVARSRATPLDLIYMA 300
DB 241 EIFFLQQAQMPBPBGMRITDTSKQNTLLSLHNAQFYLLQRTPEVARSRATPLDLIYMA 300
QY 301 LTPHPQQAQAVGVTLPSTVLPFIAGHDNTLANIGALBELNMTLPGQPDNTPPGGELVFERM 360
DB 301 LTPHPQQAQAVGVTLPSTVLPFIAGHDNTLANIGALBELNMTLPGQPDNTPPGGELVFERM 360
QY 361 RRLSDNSQWIOVSLVPQTLQOMRDKTPLSLNTPPGEVKLTLAGCEBNAQGCISLAGFTQ 420
DB 361 RRLSDNSQWIOVSLVPQTLQOMRDKTPLSLNTPPGEVKLTLAGCEBNAQGCISLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 13
Q7UD08_SHIFL PRELIMINARY, PRT, 432 AA.
AC Q7UD08;
DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2003, sequence version 1.
DE 07-MAR-2006, entry version 12.
DE Phosphatidyldiacyl phosphatase, pH 2.5 acid phosphatase.
GN Name=appa, OrderedLocustNames=S1048, ORFNames=S_1048;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA DOI=10.1128/JAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyan-Janecky L.V., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T."
RL Infect. Immun. 71:2775-2786(2003).
-----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
-----
DB EMBL: AE014073; AAP16495.1; -, Genomic_DNA.
CC
DR HSSP; P07102; 1DKL.
DR SMR; Q7UD08; 23-432.
DR GenomeReviews; AE014073.GR; S1048.
DR GO; GO:0003993; P:acid phosphatase activity; IEA.
DR InterPro; IPR000560; HisAc_phphatase.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN_1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
SQ SEQUENCE 432 AA; 47136 MW; 20044239BC6DFC CRC64;

Query Match 96.4%; Score 2218; DB 2; Length 432;
Best Local Similarity 98.6%; Pred. No. 1.5e-161;
Matches 426; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MKAILIPFLSLILPTQSAFAQSEBELKLESVVIVSRHGVRAPTKATQLMQDVTTPDAMP 60
DB 1 MKAILIPFLSLILPTQSAFAQSEBELKLESVVIVSRHGVRAPTKATQLMQDVTTPDAMP 60
QY 61 TWPVGLGMLTPRGSELIAVIGHYQQRVLVADGLAKKGCPOSQGVAAIIADVDERTRKTGE 120
DB 61 TWPVGLGMLTPRGSELIAVIGHYQQRVLVADGLAKKGCPOSQGVAAIIADVDERTRKTGE 120
QY 121 AFAAGLADCAITVHTQADTSSPDLFNPPLKTVGVCQLDNNAVTTDAIISRAGGSIAIDFTGH 180
DB 121 AFTAGLADCAITVHTQADTSSPDLFNPPLKTVGVCQLDNNAVTTDAIISRAGGSIAIDFTGH 180
```

```
QY 181 ROTAFRELERLNPQSTCLKREKQDSGSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFRELERLNPQSTCLKREKQDSGSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFFLQQAQMPBPBGMRITDTSKQNTLLSLHNAQFYLLQRTPEVARSRATPLDLIYMA 300
DB 241 EIFFLQQAQMPBPBGMRITDTSKQNTLLSLHNAQFYLLQRTPEVARSRATPLDLIYMA 300
QY 301 LTPHPQQAQAVGVTLPSTVLPFIAGHDNTLANIGALBELNMTLPGQPDNTPPGGELVFERM 360
DB 301 LTPHPQQAQAVGVTLPSTVLPFIAGHDNTLANIGALBELNMTLPGQPDNTPPGGELVFERM 360
QY 361 RRLSDNSQWIOVSLVPQTLQOMRDKTPLSLNTPPGEVKLTLAGCEBNAQGCISLAGFTQ 420
DB 361 RRLSDNSQWIOVSLVPQTLQOMRDKTPLSLNTPPGEVKLTLAGCEBNAQGCISLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 14
Q8CW75_ECOL6 PRELIMINARY, PRT, 446 AA.
AC Q8CW75;
DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2003, sequence version 1.
DE 21-FEB-2006, entry version 15.
DE Periplasmic appa protein.
GN Name=appa; ORFNames=c_1121;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O6:HI / CPT073 / ATCC 700928 / OPEC;
RX MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesech P.,
RA Raeko D., Buckles B.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Domsberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli."
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
-----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
-----
DB EMBL: AE014075; AAN79589.1; -, Genomic_DNA.
CC
DR HSSP; P07102; 1DKQ.
DR SMR; O8CW75; 33-442.
DR BioCyc; ECOL199310.C1121-MONOMER; -.
DR GO; GO:0003993; P:acid phosphatase activity; IEA.
DR InterPro; IPR000560; HisAc_phphatase.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 446 AA; 48606 MW; F1308CD691DB9F49 CRC64;

Query Match 96.1%; Score 2213; DB 2; Length 446;
Best Local Similarity 98.4%; Pred. No. 3.9e-161;
Matches 425; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 MKAILIPFLSLILPTQSAFAQSEBELKLESVVIVSRHGVRAPTKATQLMQDVTTPDAMP 60
DB 1 MKAILIPFLSLILPTQSAFAQSEBELKLESVVIVSRHGVRAPTKATQLMQDVTTPDAMP 70
QY 61 TWPVGLGMLTPRGSELIAVIGHYQQRVLVADGLAKKGCPOSQGVAAIIADVDERTRKTGE 120
DB 61 TWPVGLGMLTPRGSELIAVIGHYQQRVLVADGLAKKGCPOSQGVAAIIADVDERTRKTGE 130
```

QY 121 APAAGLAPDCAITVHTQADTSSPDPLFNPPLKTVGCQDLNANVTDAIISRAGGSTADFTGH 180
DB 131 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTVGCQDLNANVTDAIISRAGGSTADFTGH 190
QY 181 RQTAFLERLERYLNPPQSNLCIKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMT 240
DB 191 RQTAFLERLERYLNPPQSNLCIKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMT 250
QY 241 EFLILQQAQMGPEBGMGRITDSHOWNTLISLHNAQFYLLQRTPEVARSRAATPLDLIM 300
DB 251 EFLILQQAQMGPEBGMGRITDSHOWNTLISLHNAQFYLLQRTPEVARSRAATPLDLIM 310
QY 301 LTPHPPOKQAYGVTLPSTVFLIAGHDNTLANLGGALBELNMTLPQOPDNTPPGSELVFE 360
DB 311 LTPHPPOKQAYGVTLPSTVFLIAGHDNTLANLGGALBELNMTLPQOPDNTPPGSELVFE 370
QY 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPSLNTPPGEVLTLAGCERNNAQMGCSLAGFTQ 420
DB 371 RRLSDNSQMIQVSLVFQTLQOMRDKTPSLNTPPGEVLTLAGCERNNAQMGCSLAGFTQ 430
QY 421 TVNEARIPACSL 432
DB 431 TVNEARIPACSL 442

RESULT 15

Q7AFW3_ECO57
ID Q7AFW3_ECO57 PRELIMINARY; PRT; 434 AA.
AC Q7AFW3;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 10.
DE Phosphoanhydride phosphorylase.
GN OrderedLocustNames=ECS1136;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / Sakai / RIMD 050952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796; DOI=10.1093/dnares/8.1.11;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki G., Ogasawara N., Yasunaga T.,
RA Kunara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
DR EMBL; BA000007; BAB34559.1; -; Genomic_DNA.
DR SMR; Q7AFW3; 25-434.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR InterPro; IPR000560; HisAc_phosphatase.
DR Pfam; Pf00328; Acid_phosphat_A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN 1.
SQ SEQUENCE 434 AA; 47337 MW; P197DF7D1869F9C4 CRC64;

Query Match 96.0%; Score 2210; DB 2; Length 434;

Best Local Similarity 98.4%; Pred. No. 6.4e-161;

Matches 427; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 1 MKALILPFLSLILITPPOSAPAO--BPELKLESVIVSRHGVRAPTKATQLMQDVTTPDA 58
DB 1 MKALILPFLSLILITPPOSAPAOBPELKLESVIVSRHGVRAPTKATQLMQDVTTPDA 60
QY 59 WPTWPKLGLWLTTPRGELIAYLGHYQRQLVADGLLAKKCGPQSGQVAIIVADVETRTKT 118

DB 61 WPMWPKLGLWLTTPRGELIAYLGHYQRQLVADGLLTKKCGPQSGQVAIIVADVETRTKT 120
QY 119 GEAPAGLAPDCAITVHTQADTSSPDPLFNPPLKTVGCQDLNANVTDAIISRAGGSTADFT 178
DB 121 GEAPAGLAPDCAITVHTQADTSSPDPLFNPPLKTVGCQDLNANVTDAIISRAGGSTADFT 180
QY 179 GHRQTAFLERLERYLNPPQSNLCIKREKODESCSLTQALPSELKVSADNVSLTGAVSLASM 238
DB 181 GHRQTAFLERLERYLNPPQSNLCIKREKODESCSLTQALPSELKVSADNVSLTGAVSLASM 240
QY 239 LTEIFILQQAQMGPEBGMGRITDSHOWNTLISLHNAQFYLLQRTPEVARSRAATPLDLIM 298
DB 241 LTEIFILQQAQMGPEBGMGRITDSHOWNTLISLHNAQFYLLQRTPEVARSRAATPLDLIM 300
QY 299 AALTTPHPPOKQAYGVTLPSTVFLIAGHDNTLANLGGALBELNMTLPQOPDNTPPGSELVFE 358
DB 301 AALTTPHPPOKQAYGVTLPSTVFLIAGHDNTLANLGGALBELNMTLPQOPDNTPPGSELVFE 360
QY 359 RMRRLSDNSQMIQVSLVFQTLQOMRDKTPSLNTPPGEVLTLAGCERNNAQMGCSLAGF 418
DB 361 RMRRLSDNSQMIQVSLVFQTLQOMRDKTPSLNTPPGEVLTLAGCERNNAQMGCSLAGF 420
QY 419 TVNEARIPACSL 432
DB 421 TVNEARIPACSL 434

Search completed: June 13, 2006, 10:28:32
Job time : 196.752 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2006, 10:17:14 ; Search time 136.734 Seconds
(without alignments)
1444.539 Million cell updates/sec

Title: US-10-601-319-10
Perfect score: 2258
Sequence: 1 MKAILIFLTLPLTPQSA.....CSLAGPTQIVNEARIPACSL 432

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq.8:*
1: geneseqp19808:*
2: geneseqp19908:*
3: geneseqp20008:*
4: geneseqp20018:*
5: geneseqp20028:*
6: geneseqp20038:*
7: geneseqp20038:*
8: geneseqp20048:*
9: geneseqp20058:*
10: geneseqp20068:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2258	100.0	432	8	AD050300
2	2258	100.0	432	9	AED50818
3	2258	100.0	432	10	AEE75421
4	2258	100.0	436	7	ADC87743
5	2228	98.7	430	5	AAR15808
6	2228	98.7	430	6	ADA19451
7	2182	96.6	432	3	AAB36257
8	2182	96.6	432	3	AAB36261
9	2182	96.6	432	3	AAB36262
10	2182	96.6	432	3	AAB36259
11	2182	96.6	432	3	AAB36258
12	2182	96.6	432	3	AAB36263
13	2182	96.6	432	3	AAB36260
14	2182	96.6	432	3	AAB36260
15	2182	96.6	432	5	AAB15807
16	2182	96.6	432	5	AAB15807
17	2182	96.6	432	6	ADA19450
18	2182	96.6	432	6	ADA19450
19	2182	96.6	432	7	ADC87745
20	2182	96.6	432	8	AD161629
21	2182	96.6	432	8	AD050298
22	2182	96.6	432	9	AED50820
23	2182	96.6	432	10	AEE75419

24	2179	96.5	432	10	AER15220	Aef15220 Escherich
25	2177	96.4	432	4	AAR02631	Aae02631 E. coli a
26	2177	96.4	432	8	AD161631	Ad161631 Shigella
27	2172	96.2	440	8	AD050304	Ad050304 Escherich
28	2171	96.1	432	4	AAR02634	Aae02634 E. coli a
29	2171	96.1	440	4	AAB37892	Aab37892 Escherich
30	2171	96.1	440	5	AAB22836	Aae22836 Escherich
31	2171	96.1	440	5	AAB15806	Aae15806 Escherich
32	2171	96.1	440	5	ADA19446	Ada19446 E. coli B
33	2171	96.1	440	8	AD050292	Ad050292 Escherich
34	2171	96.1	440	10	AEE75413	Aee75413 E. coli B
35	2170	96.1	432	4	AAR02635	Aae02635 E. coli a
36	2169	96.1	440	8	AD050302	Ad050302 Kengaroo
37	2168.5	96.0	431	6	AAB37853	Aae37853 PNOV4054
38	2168.5	96.0	431	6	AED46568	Aed46568 NOV9x phy
39	2168.5	96.0	437	6	AAB37854	Aae37854 PNOV4058
40	2168.5	96.0	437	9	AEC39116	Aec39116 Maize 27
41	2168.5	96.0	437	9	AED46570	Aed46570 NOV9x phy
42	2157	95.5	412	6	AAB37851	Aae37851 Maize-opt
43	2157	95.5	412	7	AAB82310	Ab82310 NOV9x phy
44	2157	95.5	432	4	AAR02632	Aae02632 E. coli a
45	2153	95.3	410	5	ABP51937	Abp51937 Phytase p

ALIGNMENTS

RESULT 1	AD050300	standard; protein; 432 AA.
ID	AD050300	
AC	AD050300;	
XX	29-JUL-2004	(first entry)
XX	Escherichia coli K12	appA phytase mutant 819PH59.
DB	Escherichia coli K12	appA phytase mutant 819PH59.
XX	appA phytase; bacteria; thermal tolerance; protease stability; foodstuff;	
KW	phytase; animal feed; fish feed; dough; baking; enzyme; mutant; mutain.	
OS	Escherichia coli; K12.	
XX	Synthetic.	
XX		
FT	Key	Location/Qualifiers
FT	Misc-difference 68	/note= "Wild-type Trp replaced with Glu"
FT	FT	
FT	Misc-difference 84	/note= "Wild-type Gln replaced with Trp"
FT	FT	
FT	Misc-difference 95	/note= "Wild-type Ala replaced with Pro"
FT	FT	
FT	Misc-difference 97	/note= "Wild-type Lys replaced with Cys"
FT	FT	
FT	Misc-difference 168	/note= "Wild-type Ser replaced with Glu"
FT	FT	
FT	Misc-difference 181	/note= "Wild-type Arg replaced with Tyr"
FT	FT	
FT	Misc-difference 226	/note= "Wild-type Asn replaced with Cys"
FT	FT	
FT	Misc-difference 277	/note= "Wild-type Tyr replaced with Asp"
FT	FT	
PN	US2004091968-A1.	
XX		
PD	13-MAY-2004.	
XX		
PF	20-JUN-2003; 2003US-00601319.	
XX		
PR	13-AUG-1997; 97US-00910798.	
PR	01-MAR-1999; 99US-00259214.	
PR	13-APR-1999; 99US-00291931.	
PR	25-MAY-1999; 99US-00318528.	
PR	25-MAY-2000; 2000US-00580515.	
PR	24-MAY-2001; 2001US-00866379.	

XX New isolated or recombinant nucleic acid encoding phytase enzymes, useful
 PT as foodstuff, and for oil degumming, producing an animal feed, delivering
 PT a phytase enzyme supplement to an animal.
 PS
 SS Claim 181; SEQ ID NO 2; 104pp; English.
 XX
 CC The present invention relates to a recombinant phytase (EC 3.1.2.8)
 CC protein and its encoding also nucleic acid molecule. Phytases of the
 CC invention is a thermostable protein. They are useful in foodstuffs, for
 CC oil degumming, producing an animal feed, delivering a phytase enzyme
 CC supplement to an animal, increasing the resistance of a phytase
 CC polypeptide to enzymatic inactivation in a digestive system of an animal
 CC and in processing of corn and sorghum kernels. The invention is also
 CC useful in gene therapy and in production of transgenic animals. The
 CC present sequence is a *Becherichia coli* phytase mutant protein which is
 CC encoded by *apra* gene.
 SQ Sequence 432 AA;
 Query Match 100.0%; Score 2258; DB 9; Length 432;
 Best Local Similarity 100.0%; Pred. No. 3.2e-219;
 Matches 432; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKAILPFLSLPLTPQSAFQSEBELKESVIVSRHGVRAPTATOLMDVTPDAMP 60
 DB 1 MKAILPFLSLPLTPQSAFQSEBELKESVIVSRHGVRAPTATOLMDVTPDAMP 60
 QY 61 TWPVKLGELTPRGELIAYLGHYWRQLVADGLPKCGSPQSGQVAIADVDETRTKGS 120
 DB 61 TWPVKLGELTPRGELIAYLGHYWRQLVADGLPKCGSPQSGQVAIADVDETRTKGS 120
 QY 121 ARAAGIAPCAITVTHQADTSSPDPLFNLKTVCCQDNNANTDAILEPAGSGIADFTGH 180
 DB 121 ARAAGIAPCAITVTHQADTSSPDPLFNLKTVCCQDNNANTDAILEPAGSGIADFTGH 180
 QY 181 YGTAFRELRLVNFPOSNLCLKREKDESCSLTQALPSELKVSADCVSLTGANVSLASMLT 240
 DB 181 YGTAFRELRLVNFPOSNLCLKREKDESCSLTQALPSELKVSADCVSLTGANVSLASMLT 240
 QY 241 EIFLLQQAQGMPEPGMGRITDSHQWNTLSLHNAQFDLLQRTPEVARSATPLLDLIKTA 300
 DB 241 EIFLLQQAQGMPEPGMGRITDSHQWNTLSLHNAQFDLLQRTPEVARSATPLLDLIKTA 300
 QY 301 LTPHPQKQAYGVTLFTSVLFTAGHDNTNANTGALNELMWTLPGQPDNTPPGGELVFERM 360
 DB 301 LTPHPQKQAYGVTLFTSVLFTAGHDNTNANTGALNELMWTLPGQPDNTPPGGELVFERM 360
 QY 361 RRLSDNSQMIQVSLVFQTLQOMKDKTSLNTPPGSVKLTLAGCBERNAQGMCSLAGFTQ 420
 DB 361 RRLSDNSQMIQVSLVFQTLQOMKDKTSLNTPPGSVKLTLAGCBERNAQGMCSLAGFTQ 420
 QY 421 IYNEARIPACSL 432
 DB 421 IYNEARIPACSL 432
 RESULT 3
 AEE75421
 ID AEE75421 standard; protein; 432 AA.
 XX
 AC AEE75421;
 XX
 DT 23-FEB-2006 (first entry)
 XX
 DE E. coli B modified *apra* phytase 819PH59 SEQ ID NO: 10.
 XX
 KM enzyme; murein; thermostable; phytase; protein stabilization;
 KM pharmaceutical; fertilizer; cramp; cell culture; osteoporosis;
 KM osteopathic; food; *apra*.
 XX
 OS *Escherichia coli* K12.
 OS Synthetic.

XX
 FH Key Location/Qualifiers
 FT Misc-difference 68 /note= "Wild type residue Trp replaced by Glu"
 FT Misc-difference 72 /note= "Encoded by CGN"
 FT Misc-difference 84 /note= "Wild type residue Gln replaced by Trp"
 FT Misc-difference 95 /note= "Wild type residue Ala replaced by Pro"
 FT Misc-difference 97 /note= "Wild type residue Lys replaced by Cys"
 FT Misc-difference 168 /note= "Wild type residue Ser replaced by Glu"
 FT Misc-difference 181 /note= "Wild type residue Arg replaced by Tyr"
 FT Misc-difference 226 /note= "Wild type residue Asn replaced by Cys"
 FT Misc-difference 277 /note= "Wild type residue Tyr replaced by Asp"
 XX
 PN US2005281792-A1.
 PD 22-DEC-2005.
 XX
 PF 01-SEP-2004; 2004US-0093115.
 XX
 PR 13-AUG-1997; 97US-00910798.
 PR 01-MAR-1999; 99US-00259214.
 PR 13-APR-1999; 99US-00291931.
 PR 25-MAY-1999; 99US-00318528.
 PR 25-MAY-2000; 2000US-00580515.
 PR 24-MAY-2001; 2001US-00866379.
 XX
 PA (SHOR/) SHORT J M.
 PA (KRET/) KRETZ K A.
 PA (GRAY/) GRAY K A.
 PA (BART/) BARTON N R.
 PA (GARR/) GARRETT J B.
 PA (ODON/) O'DONOGHUE E.
 PA (BAUM/) BAUM W.
 PA (ROBE/) ROBERTSON D E.
 PA (ZORN/) ZORNER P.
 XX
 PI Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, O'Donoghue E,
 PI Baum W, Robertson DB, Zorner P;
 XX
 DR MPI: 2006-055022/06.
 XX N-PSDB; AEE75420.
 XX
 DR Formulation useful as dietary supplement for treating, preventing or
 PT reversing osteoporosis and bone loss, and preventing muscle cramps,
 PT comprises one or more polypeptides having phytase activity.
 XX
 PS Claim 1; SEQ ID NO 10; 82pp; English.
 XX
 CC The present sequence is that a novel modified phytase derived from the
 CC *Escherichia coli* *apra* protein, with improved thermal tolerance and
 CC protease stability compared to the wild-type. The present invention
 CC relates to a novel formulation, comprising at least one polypeptide
 CC having phytase activity, where the polypeptide is a modified sequence
 CC (AEE75421) derived from the *Becherichia coli* K-12 *apra* phytase
 CC (AEE75419). The modification of the enzyme, comprising a series of
 CC mutations, improves the thermal tolerance and protease stability of the
 CC protein. The specification also claims a pharmaceutical composition, a
 CC kit, an immobilized phytase; a fertilizer or soil additive; a liquid
 CC supplement for preventing muscle cramps; a hydrating agent; a tissue
 CC culture or cell culture media; and a plant food additive, all comprising
 CC the novel phytase, and a method of reducing pollution and increasing
 CC nutrient availability in an environment or environmental sample by
 CC degrading environmental phytic acid. The novel enzyme is a dietary
 CC supplement useful for treating, preventing or reversing osteoporosis or
 CC bone loss, and preventing muscle cramps. The liquid supplement is useful

CC for preventing muscle cramps. The method is useful for reducing pollution
CC and increasing nutrient availability in an environment or environmental
CC sample by degrading environmental phytic acid, where the environment or
CC environmental sample comprises a soil or a body of water. The imobilized
CC phytase is useful in foodstuffs for improving the feeding value of
CC phytate rich ingredients.

XX Sequence 432 AA;

Query Match 100.0%; Score 2258; DB 10; Length 432;
Best Local Similarity 100.0%; Pred. No. 3.2e-219;
Matches 432; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAILIPPLSLIPLTPQSAFAQSEPELKESVIVSHRGVAPPTKATQLMQDVTDPDMP 60
DB 1 MKAILIPPLSLIPLTPQSAFAQSEPELKESVIVSHRGVAPPTKATQLMQDVTDPDMP 60
QY 61 TWPVKLGELTPRGSELIVLGHYWRQRLVADGLPKCCGPOSGOVAIADVDERTKTGE 120
DB 61 TWPVKLGELTPRGSELIVLGHYWRQRLVADGLPKCCGPOSGOVAIADVDERTKTGE 120
QY 121 APAAGLADCAITVHTQADTSSPDPLFNPDKTGVQQLNANVTDAIIRAGSITADFTGH 180
DB 121 APAAGLADCAITVHTQADTSSPDPLFNPDKTGVQQLNANVTDAIIRAGSITADFTGH 180
QY 181 YOTAFRELERVUNFPQSNLCLKREKODESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
DB 181 YOTAFRELERVUNFPQSNLCLKREKODESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
QY 241 EIFLLQQAQGMPEPBGWGRITDISHQWNTLLSLHNAQFDLLQRTPEVARSRATPLDLIKTA 300
DB 241 EIFLLQQAQGMPEPBGWGRITDISHQWNTLLSLHNAQFDLLQRTPEVARSRATPLDLIKTA 300
QY 301 LTPHPPOQAYGVTLPTSVLFTAGHDNTLANLGALELNMWTLPGQPDNTPPGSELVFERW 360
DB 301 LTPHPPOQAYGVTLPTSVLFTAGHDNTLANLGALELNMWTLPGQPDNTPPGSELVFERW 360
QY 361 RRLSNSQMIQVSLVFQTLQQRDKTPLSLNTPPGSEKTLTLAGCERNAQGCISLAGFTQ 420
DB 361 RRLSNSQMIQVSLVFQTLQQRDKTPLSLNTPPGSEKTLTLAGCERNAQGCISLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 4
ADC87743
ID ADC87743 standard; protein; 436 AA.

XX AC ADC87743;
XX 01-JAN-2004 (first entry)
XX DE Modified Escherichia coli phytase.
XX KM Phytase; food supplement; enzyme delivery matrix; soybean meal;
XX KM myo-inositol-hexaphosphate; inositol; inorganic phosphate;
XX KM thermocolerant; feed value; digestion; enzyme.
XX OS Synthetic.
XX OS Escherichia coli.
XX PN US2003103958-A1.
XX PD 05-JUN-2003.
XX PF 24-MAY-2002; 2002US-00156660.
XX PR 13-AUG-1997; 97US-00910798.
XX PR 01-MAR-1999; 99US-00259214.
XX PR 13-APR-1999; 99US-00291931.

PR 25-MAY-1999; 99US-00318528.
PR 25-MAY-2000; 2000US-00580515.
PR 24-MAY-2001; 2001US-00866379.
XX (DIVER-) DIVERSA CORP.

XX Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI Machur EJ;

DR MPI; 2003-787039/74.
DR N-PSDB; ADC87742.

PT New nucleic acid encoding a polypeptide having phytase activity, useful
PT in improving the feeding value of phytate rich ingredients or as an aid
PT in phytate digestion.

PS Claim 60; SEQ ID NO 2; 113pp; English.

CC The invention discloses a new isolated or recombinant nucleic acid which
CC encodes a polypeptide having a phytase activity. Also claimed is a
CC nucleic acid probe, an amplification primer sequence pair, an expression
CC cassette comprising the nucleic acid, a vector comprising the nucleic
CC acid, a transgenic non-human animal or plant, or its seed, comprising the
CC nucleic acid, an antisense oligonucleotide, inhibiting the translation of
CC a phytase message in a cell, a heterodimer comprising the polypeptide and
CC a second domain, an array comprising immobilised polypeptide or nucleic
CC acid, a hybridoma comprising an antibody that specifically binds to the
CC polypeptide, a food supplement for an animal, an edible enzyme delivery
CC matrix, an edible pellet comprising a granule edible carrier and the
CC polypeptide, a feed composition, a soybean meal, isolating or identifying
CC the polypeptide, making an anti-phytase antibody, producing a recombinant
CC polypeptide, determining whether a compound binds to the polypeptide,
CC identifying a modulator, whole cell engineering of new or modified
CC phenotypes by using real-time metabolic flux analysis, increasing
CC thermotolerance or thermostability of the phytase polypeptide, increasing
CC the resistance of the phytase polypeptide to enzymatic inactivation in a
CC digestive system and processing of corn and sorghum kernels. The phytase
CC activity comprises catalysts of phytate (myo-inositol-hexaphosphate) to
CC inositol and inorganic phosphate and its thermocolerant. The nucleic acid
CC is useful in improving the feeding value of phytate rich ingredients or
CC as an aid in phytate digestion. The sequence presented is the modified
CC Escherichia coli phytase.

XX Sequence 436 AA;

Query Match 100.0%; Score 2258; DB 7; Length 436;
Best Local Similarity 100.0%; Pred. No. 3.3e-219;
Matches 432; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAILIPPLSLIPLTPQSAFAQSEPELKESVIVSHRGVAPPTKATQLMQDVTDPDMP 60
DB 1 MKAILIPPLSLIPLTPQSAFAQSEPELKESVIVSHRGVAPPTKATQLMQDVTDPDMP 60
QY 61 TWPVKLGELTPRGSELIVLGHYWRQRLVADGLPKCCGPOSGOVAIADVDERTKTGE 120
DB 61 TWPVKLGELTPRGSELIVLGHYWRQRLVADGLPKCCGPOSGOVAIADVDERTKTGE 120
QY 121 APAAGLADCAITVHTQADTSSPDPLFNPDKTGVQQLNANVTDAIIRAGSITADFTGH 180
DB 121 APAAGLADCAITVHTQADTSSPDPLFNPDKTGVQQLNANVTDAIIRAGSITADFTGH 180
QY 181 YOTAFRELERVUNFPQSNLCLKREKODESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
DB 181 YOTAFRELERVUNFPQSNLCLKREKODESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
QY 241 EIFLLQQAQGMPEPBGWGRITDISHQWNTLLSLHNAQFDLLQRTPEVARSRATPLDLIKTA 300
DB 241 EIFLLQQAQGMPEPBGWGRITDISHQWNTLLSLHNAQFDLLQRTPEVARSRATPLDLIKTA 300
QY 301 LTPHPPOQAYGVTLPTSVLFTAGHDNTLANLGALELNMWTLPGQPDNTPPGSELVFERW 360
DB 301 LTPHPPOQAYGVTLPTSVLFTAGHDNTLANLGALELNMWTLPGQPDNTPPGSELVFERW 360

also used for generating recombinant digestive system life forms, for producing or manufacturing alcoholic and non-alcoholic drinks based on the use of moulds, grains and/or plants, in biopulping and bio-bleaching where a reduction in the use of environmentally harmful chemicals that are traditionally used in the pulp and paper industry is desired and in the reduction or possible elimination of the need for mineral supplements, enzymes or therapeutic drugs for animals from the daily feed; thus increasing the amount calories and nutrients present in the feed.

The present sequence is *E. coli* appa phytase mutant protein

Sequence 430 AA;

Query Match 98.7%; Score 2228; DB 5; Length 430;
Best Local Similarity 99.5%; Pred. No. 3.5e-216;
Matches 430; Conservative 0; Mismatches 0; Indels 2; Gaps 2

Seq	Sequence 430 AA;	Query Match	Best Local Similarity	Matches	Conservative	Mismatches	Indels	Gaps
QY	1 MKALIIPTLSLIIPTPOSAPFQSSPELKLBSVIVSRHGVAFTKATOLMODVTPDAMP 60	1	98.7%	430	0	0	2	2
DB	1 MKALIIPTLSLIIPTPOSAPFQSSPELKLBSVIVSRHGVAFTKATOLMODVTPDAMP 60	1	99.5%	430	0	0	2	2
QY	61 TWPVYDGLTFRGSELIVIGHYWMRQRLVADGLPKCGGPOSGQVAIIVDVBERTKTGE 120	61	98.7%	430	0	0	2	2
DB	61 TWPVYDGLTFRGSELIVIGHYWMRQRLVADGLPKCGGPOSGQVAIIVDVBERTKTGE 120	61	99.5%	430	0	0	2	2
QY	121 AFAAGLAPDCAITVTTQADTSSPDDLFPNPKYGVCOLDNNATDAILERRAGSIADFTGH 180	121	98.7%	430	0	0	2	2
DB	121 AFAAGLAPDCAITVTTQADTSSPDDLFPNPKYGVCOLDNNATDAILERRAGSIADFTGH 180	121	99.5%	430	0	0	2	2
QY	181 YQTAFLERLRYLNPQSNMCLKREKODESCSLTQALPSELKVSADCVSLTGAIVLSASMLT 240	181	98.7%	430	0	0	2	2
DB	180 YQTAFLERLRYLNPQSNMCLKREKODESCSLTQALPSELKVSADCVSLTGAIVLSASMLT 239	180	99.5%	430	0	0	2	2
QY	241 EIFLLQQAQMGPEPQWGRITDTHSHONTLLSLNAQFDLQRPPEVARSATPLDLIKTA 300	241	98.7%	430	0	0	2	2
DB	240 EIFLLQQAQMGPEPQWGRITDTHSHONTLLSLNAQFDLQRPPEVARSATPLDLIKTA 298	240	99.5%	430	0	0	2	2
QY	301 LTPHPPOKQAGVTLPTSVLFTAGHDTMLANIGALBELNMTLPGQPDNTPPGGELVFERN 360	301	98.7%	430	0	0	2	2
DB	299 LTPHPPOKQAGVTLPTSVLFTAGHDTMLANIGALBELNMTLPGQPDNTPPGGELVFERN 358	299	99.5%	430	0	0	2	2
QY	361 RRLSDNSQMWISLVFQTLQQRDKTPLSLNTPPEGEVKLTLAGCERNNAQMGCSLAGFTQ 420	361	98.7%	430	0	0	2	2
DB	359 RRLSDNSQMWISLVFQTLQQRDKTPLSLNTPPEGEVKLTLAGCERNNAQMGCSLAGFTQ 418	359	99.5%	430	0	0	2	2
QY	421 IVNEARIPACSL 432	421	98.7%	430	0	0	2	2
DB	419 IVNEARIPACSL 430	419	99.5%	430	0	0	2	2

RESULT 6

ADA19451

ADA19451 standard; protein; 430 AA.

ADA19451;

20-NOV-2003 (first entry)

E. coli K12 phytase mutant.

Phytase; enzyme; phytate; appa gene; animal feed; inorganic phosphate; digestion enhancement; transgenic; thermal tolerance; protease stability; mutant; muteln.

Synthetic.

Escherichia coli; strain K12.

Key location/Qualifiers

Misc-difference 68 /note= "Wild-type Trp substituted by Glu"

Misc-difference 72 /note= "Encoded by CGN"

Misc-difference 84 /note= "Wild-type Gln substituted by Trp"

FT Misc-difference 95 /note= "Wild-type Ala substituted by Pro"
FT Misc-difference 97 /note= "Wild-type Lys substituted by Cys"
FT Misc-difference 168 /note= "Wild-type Ser substituted by Glu"
FT Misc-difference 180 /note= "Wild-type Arg substituted by Tyr"
FT Misc-difference 226 /note= "Wild-type Asn substituted by Cys"
FT Misc-difference 277 /note= "Wild-type Tyr substituted by Asp"
FT US2002136754-A1.
XX
XX
XX 26-SEP-2002.
XX
XX 24-MAY-2001; 2001US-00866379.
XX
XX 13-AUG-1997; 97US-00910798.
XX 01-MAR-1999; 99US-00259214.
XX 13-APR-1999; 99US-00291931.
XX 25-MAY-1999; 99US-00318528.
XX 25-MAY-2000; 2000US-00580515.
XX
XX (SHOR/) SHORT J M.
XX (KRET/) KRETZ K A.
XX (GRAY/) GRAY K A.
XX (BART/) BARTON N R.
XX (GAR/) GARRETT J B.
XX (DONO/) O' DONOGHUE E.
XX (MATH/) MATHUR E J.
XX
XX Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB;
XX O' Donoghue E, Mathur EJ;
XX WPI; 2003-040002/03.
XX
XX Isolated *Escherichia coli* polynucleotide encoding a modified phytase
XX enzyme, useful in the production of animal feed, for improving the
XX nutritional value of phytate-containing feedstuff and for enhancing
XX digestion in humans and animals.
XX
XX Claim 59; Fig 8; 62pp; English.
XX
XX The invention relates to an isolated *Escherichia coli* polynucleotide
XX encoding a phytase enzyme appearing as ADAl9450 and having amino acids
XX modification from W68E, Q84W, A95P, K97C, S168E, R180Y, N226C or Y277D.
XX Also included the *E. coli* appa gene ADAl9449 (or an oligonucleotide
XX derived from it) or its mutant sequence ADAl9452, expression vectors,
XX host cells, a method of improving nutritional value of a phytate-
XX containing feedstuff by contacting the phytate-containing feedstuff with
XX a pure phytase enzyme mutated phytase (where the phytase enzyme catalyzes
XX the liberation of inorganic phosphate from the phytate in the phytate-
XX containing feedstuff), a method to produce an animal feed containing a
XX microbial phytase (comprising culturing the plant cell, plant part or
XX plant under conditions where the nucleotide sequence is expressed and
XX converting the plant cells, plants or plant into a composition for animal
XX feed), a feed composition for animals (comprising the plant seeds, plant
XX cells, plant parts or plants in admixture with a phytate-containing
XX feedstuff), a method to treat a human or an animal able to benefit from
XX digestive enhancement by the activity of an exogenous phytase enzymes
XX comprising administering to the human or animal the plant seed, plant
XX cells, plant parts or plants of a transgenic plant which is modified to
XX contain an expression system which expresses a nucleotide sequence
XX encoding a phytase enzyme, a transgenic non-human organism whose genome
XX comprising a heterologous nucleic acid sequence encoding a polypeptide
XX having phytase activity. The phytase enzyme is useful for improving the
XX nutritional value of phytate-containing feedstuff, in the production of
XX animal feed and for enhancing digestion in humans and animals. The
XX invented method improves thermal tolerance and protease stability. It
XX also improves the feeding value of phytate rich ingredients. The present
XX sequence represents *E. coli* K12 mutant phytase.

XX SQ Sequence 430 AA;
Query Match 98.7%; Score 2228; DB 6; Length 430;
Best Local Similarity 99.5%; Pred. No. 3.5e-216;
Matches 430; Conservative 0; Mismatches 0; Indels 2; Gaps 2;
QY 1 MKAILIPPLSLILIPTPQSAFAOSEPELKESVIVSRHGRAPRTKATQLMQDVTDPDMP 60
DB 1 MKAILIPPLSLILIPTPQSAFAOSEPELKESVIVSRHGRAPRTKATQLMQDVTDPDMP 60
QY 61 TWPVNLGELTPRGSELIAVIGHYWRQRIVADGLIPKCGCPQSGOVAIIVADVERTKTGE 120
DB 61 TWPVNLGELTPRGSELIAVIGHYWRQRIVADGLIPKCGCPQSGOVAIIVADVERTKTGE 120
QY 121 APAAGLAPDCAITVHTQADTSSPDPLFNPILKTGVCOQDNNAVTAIIERAGSIADFTGH 180
DB 121 APAAGLAPDCAITVHTQADTSSPDPLFNPILKTGVCOQDNNAVTAIIERAGSIADFTGH 179
QY 181 YQTAFRELERVLANPPOSNLCKRREKODESCSLTQALPSELKVSDCVSLTGAVSLASMLT 240
DB 180 YQTAFRELERVLANPPOSNLCKRREKODESCSLTQALPSELKVSDCVSLTGAVSLASMLT 239
QY 241 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSHNAQFDILQRTPEVARSRAATPLDLIKTA 300
DB 240 EIFLLQQAQGMPEPGWGRITDSHQWNTLLSHNAQFD-LQRTPEVARSRAATPLDLIKTA 298
QY 301 LTPHPPOKQAYGVTLPSTVLFIAGHDTMLAMGLALBNMTLLPGOPDTPPGSELVPERW 360
DB 299 LTPHPPOKQAYGVTLPSTVLFIAGHDTMLAMGLALBNMTLLPGOPDTPPGSELVPERW 358
QY 361 RRLSDNSQMIQVSLVFTLQQRMDKTPLSLNTPGSEVCLTLAGEERNAQMGCSLAGFTQ 420
DB 359 RRLSDNSQMIQVSLVFTLQQRMDKTPLSLNTPGSEVCLTLAGEERNAQMGCSLAGFTQ 418
QY 421 IVNEARIPACSL 432
DB 419 IVNEARIPACSL 430
RESULT 7
AAB36257
ID AAB36257 standard; protein; 432 AA.
XX
XX AAB36257;
XX
XX 12-SEP-2003 (revised)
XX 20-FEB-2001 (first entry)
XX
XX lama2/APPA plasmid translated sequence.
XX
XX Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
XX environmental pollution; pig.
XX
XX Mus musculus.
XX *Escherichia coli*.
XX Chimeric.
XX WO200064247-A1.
XX
XX 02-NOV-2000.
XX
XX 20-APR-2000; 2000WO-CA000430.
XX
XX 23-APR-1999; 99US-0130508P.
XX
XX (UVGU-) UNIV GUELPH.
XX
XX Foreberg CW, Golovan S, Phillips JP;
XX WPI; 2000-687245/67.
XX
XX N-PSDB; AAC68294.
XX

Transgenic non-human animal for gastrointestinal tract specific expression of a protein, preferably phytase, comprises a nucleic acid sequence including a heterologous transgene construct encoding the protein.

Disclosure, Fig 5, 152pp; English.

The present invention provides transgenic animals which produce desired proteins, in this case pigs which express phytase in the salivary gland. Low phytase production levels result in phytate in the diet being excreted and causing phosphorus contamination in water, as well as reducing the growth of animals. The invention provides a number of transgenes containing the E. coli APPA phytase coding sequence. (Updated on 12-SEP-2003 to standardise OS field)

Sequence 432 AA;

Query Match 96.6%; Score 2182; DB 3; Length 432;
Best Local Similarity 98.1%; Pred. No. 1.6e-211;
Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 MKAIIIPFLSLIPLTPGSAFAQSEBELKLSVIVSRHGVAPATKATQMODVTPDAMP 60
DB 1 MKAIIIPFLSLIPLTPGSAFAQSEBELKLSVIVSRHGVAPATKATQMODVTPDAMP 60
QY 61 TWPVKLGELTPRGSELIVLGHYWRQRLVADGLPKCCGPOSGQVAIADVDERTKTGE 120
DB 61 TWPVKLGELTPRGSELIVLGHYWRQRLVADGLPKCCGPOSGQVAIADVDERTKTGE 120
QY 121 AFAAGLADPCAITVHTQADTSSPDLFNPPLKTGVCOLDNANTDAILSRAGSIADFTGH 180
DB 121 AFAAGLADPCAITVHTQADTSSPDLFNPPLKTGVCOLDNANTDAILSRAGSIADFTGH 180
QY 181 YOTARELERVLPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
DB 181 YOTARELERVLPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
QY 241 EIFLLQQAQGMPEPMGRITDISHQWNTLLSLHNAQFDLQRTPEVARSRATPLDLIKTA 300
DB 241 EIFLLQQAQGMPEPMGRITDISHQWNTLLSLHNAQFDLQRTPEVARSRATPLDLIKTA 300
QY 301 LTPHPQKQAVGVTLPTSVLFIAGHDTNLAMIGALBLNMTLPGQPDNTPPGSELVFERW 360
DB 301 LTPHPQKQAVGVTLPTSVLFIAGHDTNLAMIGALBLNMTLPGQPDNTPPGSELVFERW 360
QY 361 RRLSDNSQMIQVSLVFQTLQQRDKTPLSLNTPPGEVVLTLAECERNAQGCISLAGFTQ 420
DB 361 RRLSDNSQMIQVSLVFQTLQQRDKTPLSLNTPPGEVVLTLAECERNAQGCISLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 8

AAB36261 standard; protein; 432 AA.

AAB36261;

12-SEP-2003 (revised)
20-FEB-2001 (first entry)

R15/APPA plasmid translated sequence.

Transgenic animal; salivary protein; phytase; phosphorus; animal growth; environmental pollution; pig.

Rattus sp.
Escherichia coli.
Chimeric.
MO200064247-AL.

XX 02-NOV-2000.

XX 20-APR-2000; 2000WO-CA000430.

XX 23-APR-1999; 99US-0130508P.

XX (UYGU-) UNITV GUELPH.

XX Forberg CW, Golovan S, Phillips JP;

XX MPI: 2000-687245/67.

XX N-PSDB; AAC68298.

PT Transgenic non-human animal for gastrointestinal tract specific expression of a protein, preferably phytase, comprises a nucleic acid sequence including a heterologous transgene construct encoding the protein.

PS Disclosure, Fig 21, 152pp; English.

The present invention provides transgenic animals which produce desired proteins, in this case pigs which express phytase in the salivary gland. Low phytase production levels result in phytate in the diet being excreted and causing phosphorus contamination in water, as well as reducing the growth of animals. The invention provides a number of transgenes containing the E. coli APPA phytase coding sequence. (Updated on 12-SEP-2003 to standardise OS field)

Sequence 432 AA;

Query Match 96.6%; Score 2182; DB 3; Length 432;
Best Local Similarity 98.1%; Pred. No. 1.6e-211;
Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 MKAIIIPFLSLIPLTPGSAFAQSEBELKLSVIVSRHGVAPATKATQMODVTPDAMP 60
DB 1 MKAIIIPFLSLIPLTPGSAFAQSEBELKLSVIVSRHGVAPATKATQMODVTPDAMP 60
QY 61 TWPVKLGELTPRGSELIVLGHYWRQRLVADGLPKCCGPOSGQVAIADVDERTKTGE 120
DB 61 TWPVKLGELTPRGSELIVLGHYWRQRLVADGLPKCCGPOSGQVAIADVDERTKTGE 120
QY 121 AFAAGLADPCAITVHTQADTSSPDLFNPPLKTGVCOLDNANTDAILSRAGSIADFTGH 180
DB 121 AFAAGLADPCAITVHTQADTSSPDLFNPPLKTGVCOLDNANTDAILSRAGSIADFTGH 180
QY 181 YOTARELERVLPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
DB 181 YOTARELERVLPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
QY 241 EIFLLQQAQGMPEPMGRITDISHQWNTLLSLHNAQFDLQRTPEVARSRATPLDLIKTA 300
DB 241 EIFLLQQAQGMPEPMGRITDISHQWNTLLSLHNAQFDLQRTPEVARSRATPLDLIKTA 300
QY 301 LTPHPQKQAVGVTLPTSVLFIAGHDTNLAMIGALBLNMTLPGQPDNTPPGSELVFERW 360
DB 301 LTPHPQKQAVGVTLPTSVLFIAGHDTNLAMIGALBLNMTLPGQPDNTPPGSELVFERW 360
QY 361 RRLSDNSQMIQVSLVFQTLQQRDKTPLSLNTPPGEVVLTLAECERNAQGCISLAGFTQ 420
DB 361 RRLSDNSQMIQVSLVFQTLQQRDKTPLSLNTPPGEVVLTLAECERNAQGCISLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 9

AAB36262 standard; protein; 432 AA.

AAB36262;

```

XX 12-SEP-2003 (revised)
DT 20-FEB-2001 (first entry)
XX
DE SV40/APPA plasmid translated sequence.
XX Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KM environmental pollution; pig.
XX
OS Simian virus 40.
OS Escherichia coli.
OS Chimeric.
XX
PN WO200064247-A1.
XX
PD 02-NOV-2000.
XX
PF 20-APR-2000; 2000WO-CA000430.
XX
PR 23-APR-1999; 99US-0130508P.
XX
PA (UYGU-) UNIV GUELPH.
XX
PI Forsberg CM, Golovan S, Phillips JP;
XX
DR WPI; 2000-687245/67.
XX
DR N-PSDB; AAC68296.
XX
PT Transgenic non-human animal for gastrointestinal tract specific
PT expression of a protein, preferably phytase, comprises a nucleic acid
PT sequence including a heterologous transgene construct encoding the
PT protein.
XX
PS Disclosure; Fig 22; 152pp; English.
XX
CC The present invention provides transgenic animals which produce desired
CC proteins, in this case pigs which expresses phytase in the salivary
CC gland. Low phytase production levels result in phytate in the diet being
CC excreted and causing phosphorus contamination in water, as well as
CC reducing the growth of animals. The invention provides a number of
CC transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC on 12-SEP-2003 to standardise OS field)
XX
SQ Sequence 432 AA;
Query Match 96.6%; Score 2182; DB 3; Length 432;
Best Local Similarity 98.1%; Pred. No. 1.6e-211;
Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 MKAILIPLSLILPLTPOSAPFQSEPELKESVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
DB 1 MKAILIPLSLILPLTPOSAPFQSEPELKESVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
QY 61 TWPVKGELTRGGEELIAYLGHWQRVLVADGLPKCGCPQSGQVAIIADVDETRKTGE 120
DB 61 TWPVKGELTRGGEELIAYLGHWQRVLVADGLPKCGCPQSGQVAIIADVDETRKTGE 120
QY 121 AFAAGLADCAITVHTQADTSSPDLFNPPLKTVGQQLDNNAVTDAILRAGGSIAIDFTGH 180
DB 121 AFAAGLADCAITVHTQADTSSPDLFNPPLKTVGQQLDNNAVTDAILRAGGSIAIDFTGH 180
QY 181 YQTAFFRELERVYNFQSNCLKREKQDSSCSLTQALPELKVSAQCVSLTGAVSLASMLT 240
DB 181 YQTAFFRELERVYNFQSNCLKREKQDSSCSLTQALPELKVSAQCVSLTGAVSLASMLT 240
QY 241 EIFLLQQAQMPBPQGRITDSSHQWNTLLSLHNAQFDLLQRTPEVARSRAATPLDLITXA 300
DB 241 EIFLLQQAQMPBPQGRITDSSHQWNTLLSLHNAQFDLLQRTPEVARSRAATPLDLITXA 300
QY 301 LTPHPQQAQVAVTLPISVLPFAGHDITLANIGALELAWTLPQGPDPNTPPGGELVFERW 360
DB 301 LTPHPQQAQVAVTLPISVLPFAGHDITLANIGALELAWTLPQGPDPNTPPGGELVFERW 360

```

```

QY 361 RRLSDNSQWIOVSLVPOTLOQMRDPTLSLNTPGGEVYKTLAGEBENNAQMCSLAGFTQ 420
DB 361 RRLSDNSQWIOVSLVPOTLOQMRDPTLSLNTPGGEVYKTLAGEBENNAQMCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432
RESULT 10
AAB36259
ID AAB36259 standard; protein; 432 AA.
XX
AC AAB36259;
XX
DT 12-SEP-2003 (revised)
DT 20-FEB-2001 (first entry)
XX
DE R15/APPA plasmid translated sequence.
XX
KM Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KM environmental pollution; pig.
XX
OS Rattus sp.
OS Escherichia coli.
OS Chimeric.
XX
PN WO200064247-A1.
XX
PD 02-NOV-2000.
XX
PF 20-APR-2000; 2000WO-CA000430.
XX
PR 23-APR-1999; 99US-0130508P.
XX
PA (UYGU-) UNIV GUELPH.
XX
PI Forsberg CM, Golovan S, Phillips JP;
XX
DR WPI; 2000-687245/67.
XX
DR N-PSDB; AAC68296.
XX
PT Transgenic non-human animal for gastrointestinal tract specific
PT expression of a protein, preferably phytase, comprises a nucleic acid
PT sequence including a heterologous transgene construct encoding the
PT protein.
XX
PS Disclosure; Fig 19; 152pp; English.
XX
CC The present invention provides transgenic animals which produce desired
CC proteins, in this case pigs which expresses phytase in the salivary
CC gland. Low phytase production levels result in phytate in the diet being
CC excreted and causing phosphorus contamination in water, as well as
CC reducing the growth of animals. The invention provides a number of
CC transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC on 12-SEP-2003 to standardise OS field)
XX
SQ Sequence 432 AA;
Query Match 96.6%; Score 2182; DB 3; Length 432;
Best Local Similarity 98.1%; Pred. No. 1.6e-211;
Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 MKAILIPLSLILPLTPOSAPFQSEPELKESVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
DB 1 MKAILIPLSLILPLTPOSAPFQSEPELKESVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
QY 61 TWPVKGELTRGGEELIAYLGHWQRVLVADGLPKCGCPQSGQVAIIADVDETRKTGE 120
DB 61 TWPVKGELTRGGEELIAYLGHWQRVLVADGLPKCGCPQSGQVAIIADVDETRKTGE 120
QY 121 AFAAGLADCAITVHTQADTSSPDLFNPPLKTVGQQLDNNAVTDAILRAGGSIAIDFTGH 180

```

Db 121 APAAGLAPCAITVTHQADTSPPDLFNPDKTGVCOLDANVTDAILSRAGSIADFTGH 180
 QY 181 YQTAFFRELERVLPFQSNLCLKREKODESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
 Db 181 RQTAFFRELERVLPFQSNLCLKREKODESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
 QY 241 EIFFLQQAQMGPEPGMGRIITDSHQNNTLLSLNNAQFDLQRTPEVARSRATPLDLIKTA 300
 Db 241 EIFFLQQAQMGPEPGMGRIITDSHQNNTLLSLNNAQFDLQRTPEVARSRATPLDLIKTA 300
 QY 301 LTPHPPOKQAYGVTLPTSVLFIAGHDTNLANTGALBELNMTLPQGPDNTPPGSELVFERM 360
 Db 301 LTPHPPOKQAYGVTLPTSVLFIAGHDTNLANTGALBELNMTLPQGPDNTPPGSELVFERM 360
 QY 361 RLSDNSQMIQVSLVFQTLQQRDKTPLSLNTPPGEVKLTLAGCERNNAQMGCSLAGFTQ 420
 Db 361 RLSDNSQMIQVSLVFQTLQQRDKTPLSLNTPPGEVKLTLAGCERNNAQMGCSLAGFTQ 420
 QY 421 IYNEARIPACSL 432
 Db 421 IYNEARIPACSL 432

RESULT 11

AAB36258
 ID AAB36258 standard; protein; 432 AA.
 XX AAB36258;
 AC AAB36258;
 XX 12-SEP-2003 (revised)
 DT 20-FEB-2001 (first entry)
 XX
 DE R15/APPA plasmid translated sequence.
 XX
 KM Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
 KM environmental pollution; pig.
 XX
 OS Rattus sp.
 OS Escherichia coli.
 OS Chimeric.
 XX
 PN WO200064247-A1.
 XX
 PD 02-NOV-2000.
 XX
 PF 20-APR-2000; 2000WO-CA000430.
 XX
 PR 23-APR-1999; 99US-0130508P.
 XX
 PA (UYGU-) UNIV GUEBLPH.
 XX
 PI Foreberg CW, Golovan S, Phillips JP;
 XX
 DR WPI, 2000-687245/67.
 DR N-PSDB; AAC68295.
 XX
 PT Transgenic non-human animal for gastrointestinal tract specific
 PT expression of a protein, preferably phytase, comprises a nucleic acid
 PT sequence including a heterologous transgene construct encoding the
 PT protein.
 XX
 PS Disclosure; Fig 18; 152pp; English.
 XX
 CC The present invention provides transgenic animals which produce desired
 CC proteins, in this case pigs which expresses phytase in the salivary
 CC gland. Low phytase production levels result in phytate in the diet being
 CC excreted and causing phosphorus contamination in water, as well as
 CC reducing the growth of animals. The invention provides a number of
 CC transgenes containing the E. coli APPA phytase coding sequence. (Updated
 CC on 12-SEP-2003 to standardise OS field)
 XX
 XX Sequence 432 AA;
 SQ

Query Match 96.6%; Score 2182; DB 3; Length 432;
 Best Local Similarity 98.1%; Pred. No. 1.6e-211;
 Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 MKAIIIPFLSLIPLTPQSAFAQSEBELKESVIVSRGVAPATKATQMLQDVTDPDAMP 60
 Db 1 MKAIIIPFLSLIPLTPQSAFAQSEBELKESVIVSRGVAPATKATQMLQDVTDPDAMP 60
 QY 61 TWPVKGELTPRGSELINLVGHYRQRLVADGLKRCGPGSGOVAIIVDVERTRKGE 120
 Db 61 TWPVKGELTPRGSELINLVGHYRQRLVADGLKRCGPGSGOVAIIVDVERTRKGE 120
 QY 121 APAAGLAPCAITVTHQADTSPPDLFNPDKTGVCOLDANVTDAILSRAGSIADFTGH 180
 Db 121 APAAGLAPCAITVTHQADTSPPDLFNPDKTGVCOLDANVTDAILSRAGSIADFTGH 180
 QY 181 YQTAFFRELERVLPFQSNLCLKREKODESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
 Db 181 YQTAFFRELERVLPFQSNLCLKREKODESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
 QY 241 EIFFLQQAQMGPEPGMGRIITDSHQNNTLLSLNNAQFDLQRTPEVARSRATPLDLIKTA 300
 Db 241 EIFFLQQAQMGPEPGMGRIITDSHQNNTLLSLNNAQFDLQRTPEVARSRATPLDLIKTA 300
 QY 301 LTPHPPOKQAYGVTLPTSVLFIAGHDTNLANTGALBELNMTLPQGPDNTPPGSELVFERM 360
 Db 301 LTPHPPOKQAYGVTLPTSVLFIAGHDTNLANTGALBELNMTLPQGPDNTPPGSELVFERM 360
 QY 361 RLSDNSQMIQVSLVFQTLQQRDKTPLSLNTPPGEVKLTLAGCERNNAQMGCSLAGFTQ 420
 Db 361 RLSDNSQMIQVSLVFQTLQQRDKTPLSLNTPPGEVKLTLAGCERNNAQMGCSLAGFTQ 420
 QY 421 IYNEARIPACSL 432
 Db 421 IYNEARIPACSL 432

RESULT 12

AAB36263
 ID AAB36263 standard; protein; 432 AA.
 XX AAB36263;
 AC AAB36263;
 XX 12-SEP-2003 (revised)
 DT 20-FEB-2001 (first entry)
 XX
 DE Lama2/APPA plasmid translated sequence.
 XX
 KM Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
 KM environmental pollution; pig.
 XX
 OS Mus musculus.
 OS Escherichia coli.
 OS Chimeric.
 XX
 PN WO200064247-A1.
 XX
 PD 02-NOV-2000.
 XX
 PF 20-APR-2000; 2000WO-CA000430.
 XX
 PR 23-APR-1999; 99US-0130508P.
 XX
 PA (UYGU-) UNIV GUEBLPH.
 XX
 PI Foreberg CW, Golovan S, Phillips JP;
 XX
 DR WPI, 2000-687245/67.
 DR N-PSDB; AAC68300.
 XX
 PT Transgenic non-human animal for gastrointestinal tract specific
 PT expression of a protein, preferably phytase, comprises a nucleic acid
 PT sequence including a heterologous transgene construct encoding the

PT protein.

PS Disclosure; Fig 23, 152bp; English.

XX The present invention provides transgenic animals which produce desired
XX proteins, in this case pigs which expresses phytase in the salivary
CC gland. Low phytase production levels result in phytate in the diet being
CC excreted and causing phosphorus contamination in water, as well as
CC reducing the growth of animals. The invention provides a number of
CC transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC on 12-SEP-2003 to standardise OS field)

XX Sequence 432 AA;

Query Match 96.6%; Score 2182; DB 3; Length 432;
Best Local Similarity 98.1%; Pred. No. 1.6e-211;
Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 MKAILIPLSLILPTPOSAPFAOSEPELKESVIVSHGVRAPTKATQLMQDVTDPDAMP 60
DB 1 MKAILIPLSLILPTPOSAPFAOSEPELKESVIVSHGVRAPTKATQLMQDVTDPDAMP 60
QY 61 TWPVKLGELTPRGSELIAVLGHYWRQRLVADGLPKCGCPQSGQVAIIADVERTRTKGE 120
DB 61 TWPVKLGELTPRGSELIAVLGHYWRQRLVADGLPKCGCPQSGQVAIIADVERTRTKGE 120
QY 121 AFAAGLADCAITVHTQADTSSPDPLFNPPLKTGVQQLDNANVTDAIISRAGGSIAADFTGH 180
DB 121 AFAAGLADCAITVHTQADTSSPDPLFNPPLKTGVQQLDNANVTDAIISRAGGSIAADFTGH 180
QY 181 YOTAFRELERVLNFPQSNLCKREKQDSCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
DB 181 ROTAFRELERVLNFPQSNLCKREKQDSCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
QY 241 EIFLLQQAQGMPEPQWGRITDSSHQWNTLLSLHNAQFDLLQRTPEVARSRAATPLDLIKTA 300
DB 241 EIFLLQQAQGMPEPQWGRITDSSHQWNTLLSLHNAQFDLLQRTPEVARSRAATPLDLIKTA 300
QY 301 LTPHPPOKQAYGVTLPTSVLPIAGHDITNLAMLGALBELNMTLPQGPDMTPPGGELVFERRW 360
DB 301 LTPHPPOKQAYGVTLPTSVLPIAGHDITNLAMLGALBELNMTLPQGPDMTPPGGELVFERRW 360
QY 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPPGEVGLTLAGCERNAQGMCSLAGFTQ 420
DB 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPPGEVGLTLAGCERNAQGMCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 13

AAB36260 standard; protein; 432 AA.

XX AAB36260;
XX
XX 12-SEP-2003 (revised)
DT 20-FEB-2001 (first entry)
XX
XX R15/APPA plaemid translated sequence.
DE
XX Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KM environmental pollution; pig.
XX
XX Rattus sp.
OS Escherichia coli.
OS Chimeric.
XX
XX WO200064247-A1.
XX
XX 02-NOV-2000.

PF 20-APR-2000; 2000WO-CA000430.

XX 23-APR-1999; 99US-0130508P.

XX (UYGU-) UNIV GUELPH.

PI Forsberg CM, Golovan S, Phillips JP;

XX WPI; 2000-687245/67.

DR N-PSDB; AAC68297.

XX Transgenic non-human animal for gastrointestinal tract specific
PT expression of a protein, preferably phytase, comprises a nucleic acid
PT sequence including a heterologous transgene construct encoding the
PT protein.

PS Disclosure; Fig 20, 152bp; English.

XX The present invention provides transgenic animals which produce desired
CC proteins, in this case pigs which expresses phytase in the salivary
CC gland. Low phytase production levels result in phytate in the diet being
CC excreted and causing phosphorus contamination in water, as well as
CC reducing the growth of animals. The invention provides a number of
CC transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC on 12-SEP-2003 to standardise OS field)

XX Sequence 432 AA;

Query Match 96.6%; Score 2182; DB 3; Length 432;
Best Local Similarity 98.1%; Pred. No. 1.6e-211;
Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 MKAILIPLSLILPTPOSAPFAOSEPELKESVIVSHGVRAPTKATQLMQDVTDPDAMP 60
DB 1 MKAILIPLSLILPTPOSAPFAOSEPELKESVIVSHGVRAPTKATQLMQDVTDPDAMP 60
QY 61 TWPVKLGELTPRGSELIAVLGHYWRQRLVADGLPKCGCPQSGQVAIIADVERTRTKGE 120
DB 61 TWPVKLGELTPRGSELIAVLGHYWRQRLVADGLPKCGCPQSGQVAIIADVERTRTKGE 120
QY 121 AFAAGLADCAITVHTQADTSSPDPLFNPPLKTGVQQLDNANVTDAIISRAGGSIAADFTGH 180
DB 121 AFAAGLADCAITVHTQADTSSPDPLFNPPLKTGVQQLDNANVTDAIISRAGGSIAADFTGH 180
QY 181 YOTAFRELERVLNFPQSNLCKREKQDSCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
DB 181 ROTAFRELERVLNFPQSNLCKREKQDSCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
QY 241 EIFLLQQAQGMPEPQWGRITDSSHQWNTLLSLHNAQFDLLQRTPEVARSRAATPLDLIKTA 300
DB 241 EIFLLQQAQGMPEPQWGRITDSSHQWNTLLSLHNAQFDLLQRTPEVARSRAATPLDLIKTA 300
QY 301 LTPHPPOKQAYGVTLPTSVLPIAGHDITNLAMLGALBELNMTLPQGPDMTPPGGELVFERRW 360
DB 301 LTPHPPOKQAYGVTLPTSVLPIAGHDITNLAMLGALBELNMTLPQGPDMTPPGGELVFERRW 360
QY 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPPGEVGLTLAGCERNAQGMCSLAGFTQ 420
DB 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPPGEVGLTLAGCERNAQGMCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 14

AAU77775 standard; protein; 432 AA.

XX AAU77775;
XX
XX 05-JUN-2002 (first entry)
DT
XX

Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCQOLDNANVTDAILE-AGGSIADFTG 179
Qy 181 YQTAFRELERVLANPQSNCLCKREKODESCSLTQALPSELKVSADCVSLTGAVALASMLT 240
Db 180 YQTAFRELERVLANPQSNCLCKREKODESCSLTQALPSELKVSADCVSLTGAVALASMLT 239
Qy 241 EIFLLQQAQGMPEBGMGRITDSHOMNTLLSLHNAQFDLQRTPEVARSRAATPDLILKTA 300
Db 240 EIFLLQQAQGMPEBGMGRITDSHOMNTLLSLHNAQFD-LQRTPEVARSRAATPDLILKTA 298
Qy 301 LTPHPQKQAYGVTLPTSVLPIAGHDNTLANLGGALBELNMTLLPGQPDNTPPGGELVFERW 360
Db 299 LTPHPQKQAYGVTLPTSVLPIAGHDNTLANLGGALBELNMTLLPGQPDNTPPGGELVFERW 358
Qy 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPSLANTPPGEVKLTLAGCEBRNAQMGCSLAGFTQ 420
Db 359 RRLSDNSQWIOVSLVFQTLQOMRDKTPSLANTPPGEVKLTLAGCEBRNAQMGCSLAGFTQ 418
Qy 421 IVNEARIPACSL 432
Db 419 IVNEARIPACSL 430

RESULT 2

US-09-866-379-8
; Sequence 8, Application US/09866379
; Patent No. 6855365
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-866-379-8

Query Match 96.6%; Score 2182; DB 2; Length 432;
Best Local Similarity 98.1%; Pred. No. 1.8e-222;
Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 MKAILIPLSLILPLTPQSAFAQSEPELKESVTVSRHGVRAPTKATQOLMODVTPDAMP 60
Db 1 MKAILIPLSLILPLTPQSAFAQSEPELKESVTVSRHGVRAPTKATQOLMODVTPDAMP 60
Qy 61 TWPVKLGELTPRGSELIAVIGHYWRQRLVADGLPKCGCPQSGQVAILIADVDETRKTGE 120
Db 61 TWPVKLGELTPRGSELIAVIGHYWRQRLVADGLPKCGCPQSGQVAILIADVDETRKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCQOLDNANVTDAILERAGGSIAADFTG 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCQOLDNANVTDAILERAGGSIAADFTG 180

Qy 181 YQTAFRELERVLANPQSNCLCKREKODESCSLTQALPSELKVSADCVSLTGAVALASMLT 240
Db 181 YQTAFRELERVLANPQSNCLCKREKODESCSLTQALPSELKVSADCVSLTGAVALASMLT 240
Qy 241 EIFLLQQAQGMPEBGMGRITDSHOMNTLLSLHNAQFDLQRTPEVARSRAATPDLILKTA 300
Db 241 EIFLLQQAQGMPEBGMGRITDSHOMNTLLSLHNAQFDLQRTPEVARSRAATPDLILKTA 300
Qy 301 LTPHPQKQAYGVTLPTSVLPIAGHDNTLANLGGALBELNMTLLPGQPDNTPPGGELVFERW 360
Db 301 LTPHPQKQAYGVTLPTSVLPIAGHDNTLANLGGALBELNMTLLPGQPDNTPPGGELVFERW 360
Qy 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPSLANTPPGEVKLTLAGCEBRNAQMGCSLAGFTQ 420
Db 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPSLANTPPGEVKLTLAGCEBRNAQMGCSLAGFTQ 420
Qy 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432

RESULT 3

US-09-715-477-1
; Sequence 1, Application US/09715477
; Patent No. 6841370
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYTASE
; FILE REFERENCE: 19603/4031
; CURRENT APPLICATION NUMBER: US/09/715,477
; CURRENT FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,179
; PRIOR FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-715-477-1

Query Match 96.4%; Score 2177; DB 2; Length 432;
Best Local Similarity 97.9%; Pred. No. 6.1e-222;
Matches 423; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 MKAILIPLSLILPLTPQSAFAQSEPELKESVTVSRHGVRAPTKATQOLMODVTPDAMP 60
Db 1 MKAILIPLSLILPLTPQSAFAQSEPELKESVTVSRHGVRAPTKATQOLMODVTPDAMP 60
Qy 61 TWPVKLGELTPRGSELIAVIGHYWRQRLVADGLPKCGCPQSGQVAILIADVDETRKTGE 120
Db 61 TWPVKLGELTPRGSELIAVIGHYWRQRLVADGLPKCGCPQSGQVAILIADVDETRKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCQOLDNANVTDAILERAGGSIAADFTG 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCQOLDNANVTDAILERAGGSIAADFTG 180
Qy 181 YQTAFRELERVLANPQSNCLCKREKODESCSLTQALPSELKVSADCVSLTGAVALASMLT 240
Db 181 YQTAFRELERVLANPQSNCLCKREKODESCSLTQALPSELKVSADCVSLTGAVALASMLT 240
Qy 241 EIFLLQQAQGMPEBGMGRITDSHOMNTLLSLHNAQFDLQRTPEVARSRAATPDLILKTA 300
Db 241 EIFLLQQAQGMPEBGMGRITDSHOMNTLLSLHNAQFDLQRTPEVARSRAATPDLILKTA 300
Qy 301 LTPHPQKQAYGVTLPTSVLPIAGHDNTLANLGGALBELNMTLLPGQPDNTPPGGELVFERW 360
Db 301 LTPHPQKQAYGVTLPTSVLPIAGHDNTLANLGGALBELNMTLLPGQPDNTPPGGELVFERW 360
Qy 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPSLANTPPGEVKLTLAGCEBRNAQMGCSLAGFTQ 420
Db 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPSLANTPPGEVKLTLAGCEBRNAQMGCSLAGFTQ 420

Qy 421 IYNEARIPACSL 432
Db 421 IYNEARIPACSL 432

RESULT 4
US-09-259-214-2
; Sequence 2, Application US/09259214A
; Patent No. 6110719
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: DIVER1370-1
; CURRENT APPLICATION NUMBER: US/09/259,214A
; EARLIER FILING DATE: 1999-03-01
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-259-214-2

Query Match 96.1%; Score 2171; DB 2; Length 440;
Best Local Similarity 97.7%; Pred. No. 2.7e-221;
Matches 422; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 MKAILIPFLSLIPLTPQSAFAQSEBELKESVIVSRHGVAPPTKATQLMQDVTDPDAMP 60
Db 1 MKAILIPFLSLIPLTPQSAFAQSEBELKESVIVSRHGVAPPTKATQLMQDVTDPDAMP 60
Qy 61 TWPVKGLBTPRGSELIAVLYGHYQRQLVADGLPKCGCPQSGQVAIADVDERTKIGE 120
Db 61 TWPVKGLBTPRGSELIAVLYGHYQRQLVADGLPKCGCPQSGQVAIADVDERTKIGE 120
Qy 121 AFAAGIAPDCAITVHTQADTSSPDLFNPDKTGVQCLDNANVTDAILEPAGSIADFTGH 180
Db 121 AFAAGIAPDCAITVHTQADTSSPDLFNPDKTGVQCLDNANVTDAILEPAGSIADFTGH 180
Qy 181 YGTARELERVNFPOSNCLKREKODESCSLTQALPSELKVSADCVSLTGAVSLASMT 240
Db 181 YGTARELERVNFPOSNCLKREKODESCSLTQALPSELKVSADCVSLTGAVSLASMT 240
Qy 241 EIFLLQQAQGMPEPGMGRITDSHQNNTLLSLHNAQFDLLQRTPEVARSATPLDLIKTA 300
Db 241 EIFLLQQAQGMPEPGMGRITDSHQNNTLLSLHNAQFDLLQRTPEVARSATPLDLIKTA 300
Qy 301 LTPHPPOKQAYGVTLPTSVLFTAGHDYNIANIGALIELNMTLPQGPNDTPPGSELVFERW 360
Db 301 LTPHPPOKQAYGVTLPTSVLFTAGHDYNIANIGALIELNMTLPQGPNDTPPGSELVFERW 360
Qy 361 RLNSNSQMIQVSLVFQTLQQRDKTPLSLNTPPGSEVQLTLAGCERNAGWCISLAGFTQ 420
Db 361 RLNSNSQMIQVSLVFQTLQQRDKTPLSLNTPPGSEVQLTLAGCERNAGWCISLAGFTQ 420
Qy 421 IYNEARIPACSL 432
Db 421 IYNEARIPACSL 432

RESULT 5
US-09-318-528-2
; Sequence 2, Application US/09318528
; Patent No. 6183740
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/318,528
; CURRENT FILING DATE: 1999-05-25

; EARLIER APPLICATION NUMBER: 09/291,931
; EARLIER FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-318-528-2

Query Match 96.1%; Score 2171; DB 2; Length 440;
Best Local Similarity 97.7%; Pred. No. 2.7e-221;
Matches 422; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 MKAILIPFLSLIPLTPQSAFAQSEBELKESVIVSRHGVAPPTKATQLMQDVTDPDAMP 60
Db 1 MKAILIPFLSLIPLTPQSAFAQSEBELKESVIVSRHGVAPPTKATQLMQDVTDPDAMP 60
Qy 61 TWPVKGLBTPRGSELIAVLYGHYQRQLVADGLPKCGCPQSGQVAIADVDERTKIGE 120
Db 61 TWPVKGLBTPRGSELIAVLYGHYQRQLVADGLPKCGCPQSGQVAIADVDERTKIGE 120
Qy 121 AFAAGIAPDCAITVHTQADTSSPDLFNPDKTGVQCLDNANVTDAILEPAGSIADFTGH 180
Db 121 AFAAGIAPDCAITVHTQADTSSPDLFNPDKTGVQCLDNANVTDAILEPAGSIADFTGH 180
Qy 181 YGTARELERVNFPOSNCLKREKODESCSLTQALPSELKVSADCVSLTGAVSLASMT 240
Db 181 YGTARELERVNFPOSNCLKREKODESCSLTQALPSELKVSADCVSLTGAVSLASMT 240
Qy 241 EIFLLQQAQGMPEPGMGRITDSHQNNTLLSLHNAQFDLLQRTPEVARSATPLDLIKTA 300
Db 241 EIFLLQQAQGMPEPGMGRITDSHQNNTLLSLHNAQFDLLQRTPEVARSATPLDLIKTA 300
Qy 301 LTPHPPOKQAYGVTLPTSVLFTAGHDYNIANIGALIELNMTLPQGPNDTPPGSELVFERW 360
Db 301 LTPHPPOKQAYGVTLPTSVLFTAGHDYNIANIGALIELNMTLPQGPNDTPPGSELVFERW 360
Qy 361 RLNSNSQMIQVSLVFQTLQQRDKTPLSLNTPPGSEVQLTLAGCERNAGWCISLAGFTQ 420
Db 361 RLNSNSQMIQVSLVFQTLQQRDKTPLSLNTPPGSEVQLTLAGCERNAGWCISLAGFTQ 420
Qy 421 IYNEARIPACSL 432
Db 421 IYNEARIPACSL 432

RESULT 6
US-09-291-931-2
; Sequence 2, Application US/09291931A
; Patent No. 6190897
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/291,931A
; EARLIER FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-291-931-2

Query Match 96.1%; Score 2171; DB 2; Length 440;
Best Local Similarity 97.7%; Pred. No. 2,7e-221;
Matches 422; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 MKAILIPLSLIPLTPQSAFAQSEPEIKLESVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
DB 1 MKAILIPLSLIPLTPQSAFAQSEPEIKLESVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
QY 61 TWPVGLGMLTRPGGELIAYIGHYRQRLVADGLPKCGCPQSGQVAIIADVDETRKTGE 120
DB 61 TWPVGLGMLTRPGGELIAYIGHYRQRLVADGLPKCGCPQSGQVAIIADVDETRKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCQOLDNNVTDAILERAGGSIAIDFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCQOLDNNVTDAILERAGGSIAIDFTGH 180
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCQOLDNNVTDAILERAGGSIAIDFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCQOLDNNVTDAILERAGGSIAIDFTGH 180
QY 181 YQTAFRELERLVNFPQSNCLCKREKQDSCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
DB 181 YQTAFRELERLVNFPQSNCLCKREKQDSCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
QY 241 EIFLLOQAQMPPEGMGRITDSHQNNTLSLHNAQFYLLOKTPVARSRAITPLDLIKTA 300
DB 241 EIFLLOQAQMPPEGMGRITDSHQNNTLSLHNAQFYLLOKTPVARSRAITPLDLIKTA 300
QY 301 LTPHPQKQAVGVTLPFVSFLFIAGHDTNLNMGALBLNMTLPGQPDNTPPGGELVFERW 360
DB 301 LTPHPQKQAVGVTLPFVSFLFIAGHDTNLNMGALBLNMTLPGQPDNTPPGGELVFERW 360
QY 361 RRLSDNSQWIOVSLVFQTLQOMRDXTPLSLNTPPGEVYKTLTAGCEERNAQMCSLAGFTQ 420
DB 361 RRLSDNSQWIOVSLVFQTLQOMRDXTPLSLNTPPGEVYKTLTAGCEERNAQMCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 7
US-09-580-515-2
; Sequence 2, Application US/09580515
; Patent No. 6720014
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYLASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/580,515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-580-515-2

Query Match 96.1%; Score 2171; DB 2; Length 440;
Best Local Similarity 97.7%; Pred. No. 2,7e-221;
Matches 422; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 MKAILIPLSLIPLTPQSAFAQSEPEIKLESVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
DB 1 MKAILIPLSLIPLTPQSAFAQSEPEIKLESVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
QY 61 TWPVGLGMLTRPGGELIAYIGHYRQRLVADGLPKCGCPQSGQVAIIADVDETRKTGE 120
DB 61 TWPVGLGMLTRPGGELIAYIGHYRQRLVADGLPKCGCPQSGQVAIIADVDETRKTGE 120

DB 61 TWPVGLGMLTRPGGELIAYIGHYRQRLVADGLPKCGCPQSGQVAIIADVDETRKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCQOLDNNVTDAILERAGGSIAIDFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCQOLDNNVTDAILERAGGSIAIDFTGH 180
QY 181 YQTAFRELERLVNFPQSNCLCKREKQDSCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
DB 181 YQTAFRELERLVNFPQSNCLCKREKQDSCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
QY 241 EIFLLOQAQMPPEGMGRITDSHQNNTLSLHNAQFYLLOKTPVARSRAITPLDLIKTA 300
DB 241 EIFLLOQAQMPPEGMGRITDSHQNNTLSLHNAQFYLLOKTPVARSRAITPLDLIKTA 300
QY 301 LTPHPQKQAVGVTLPFVSFLFIAGHDTNLNMGALBLNMTLPGQPDNTPPGGELVFERW 360
DB 301 LTPHPQKQAVGVTLPFVSFLFIAGHDTNLNMGALBLNMTLPGQPDNTPPGGELVFERW 360
QY 361 RRLSDNSQWIOVSLVFQTLQOMRDXTPLSLNTPPGEVYKTLTAGCEERNAQMCSLAGFTQ 420
DB 361 RRLSDNSQWIOVSLVFQTLQOMRDXTPLSLNTPPGEVYKTLTAGCEERNAQMCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 8
US-09-866-379-2
; Sequence 2, Application US/09866379
; Patent No. 6853365
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Bileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYLASES AND USES THEREOF
; FILE REFERENCE: DIVERSA370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-866-379-2

Query Match 96.1%; Score 2171; DB 2; Length 440;
Best Local Similarity 97.7%; Pred. No. 2,7e-221;
Matches 422; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 MKAILIPLSLIPLTPQSAFAQSEPEIKLESVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
DB 1 MKAILIPLSLIPLTPQSAFAQSEPEIKLESVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
QY 61 TWPVGLGMLTRPGGELIAYIGHYRQRLVADGLPKCGCPQSGQVAIIADVDETRKTGE 120
DB 61 TWPVGLGMLTRPGGELIAYIGHYRQRLVADGLPKCGCPQSGQVAIIADVDETRKTGE 120

```

QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOLDNANVTDAIILERAGGSIAIDFTGH 180
|
|
|
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOLDNANVTDAIILSRAGGSIAIDFTGH 180
|
|
|
QY 181 YOTAFRELERVINFPOSNTCLKREKODESCSLTQALPSELKVSADCVSLTGAIVSLASMLT 240
|
|
|
DB 181 ROTAFRELERVINFPOSNTCLKREKODESCSLTQALPSELKVSADNVSLTGAIVSLASMLT 240
|
|
|
QY 241 EYFLLOAOAGMPBPGRITDTSQWNTLLSLHNAQFDLLQRTPEVARSRATPLDLIKTA 300
|
|
|
DB 241 EYFLLOAOAGMPBPGRITDTSQWNTLLSLHNAQFDLLQRTPEVARSRATPLDLIKTA 300
|
|
|
QY 301 LTPHPPOKQAYGVTLPTSVLFIAGHDNTLANIGALIELNMTLPGQPDNTPPGSELVFERW 360
|
|
|
DB 301 LTPHPPOKQAYGVTLPTSVLFIAGHDNTLANIGALIELNMTLPGQPDNTPPGSELVFERW 360
|
|
|
QY 361 RRLSDNSQWIOVSLVFQTLQQRMDKTPLSLNTPPGEVKTLLAGCERRNAQMGCSLAGFTQ 420
|
|
|
DB 361 RRLSDNSQWIOVSLVFQTLQQRMDKTPLSLNTPPGEVKTLLAGCERRNAQMGCSLAGFTQ 420
|
|
|
QY 421 IVNEARIPACSL 432
|
|
|
DB 421 IVNEARIPACSL 432
|
|
|

```

RESULT 9

```

US-09-715-477-3
; Sequence 3, Application US/09715477
; Patent No. 6841370
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYTASE
; FILE REFERENCE: 19603/4031
; CURRENT APPLICATION NUMBER: US/09/715,477
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,179
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 432
; TYPE: PRF
; ORGANISM: Escherichia coli
US-09-715-477-3

```

```

Query Match 95.5%; Score 2157; DB 2; Length 432;
Best Local Similarity 97.2%; Pred. No. 8e-220;
Matches 420; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

```

```

QY 1 MAAILIPFLSLIPLTPQSAFAQSEBELKLESVIVSRHGVAPPTKATQMLQDVTDPDAMP 60
|
|
|
DB 1 MAAILIPFLSLIPLTPQSAFAQSEBELKLESVIVSRHGVAPPTKATQMLQDVTDPDAMP 60
|
|
|
QY 61 TWPVGLGELTPRGSELIAVLGHYWRQLVADGLPKCCGPOSGQVAIADVDERTRKTGE 120
|
|
|
DB 61 TWPVGLGELTPRGSELIAVLGHYWRQLVADGLPKCCGPOSGQVAIADVDERTRKTGE 120
|
|
|
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOLDNANVTDAIILERAGGSIAIDFTGH 180
|
|
|
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOLDNANVTDAIILSRAGGSIAIDFTGH 180
|
|
|
QY 181 YOTAFRELERVINFPOSNTCLKREKODESCSLTQALPSELKVSADCVSLTGAIVSLASMLT 240
|
|
|
DB 181 ROTAFRELERVINFPOSNTCLKREKODESCSLTQALPSELKVSADNVSLTGAIVSLASMLT 240
|
|
|
QY 241 EYFLLOAOAGMPBPGRITDTSQWNTLLSLHNAQFDLLQRTPEVARSRATPLDLIKTA 300
|
|
|
DB 241 EYFLLOAOAGMPBPGRITDTSQWNTLLSLHNAQFDLLQRTPEVARSRATPLDLIKTA 300
|
|
|
QY 301 LTPHPPOKQAYGVTLPTSVLFIAGHDNTLANIGALIELNMTLPGQPDNTPPGSELVFERW 360
|
|
|
DB 301 LTPHPPOKQAYGVTLPTSVLFIAGHDNTLANIGALIELNMTLPGQPDNTPPGSELVFERW 360
|
|
|

```

```

QY 361 RRLSDNSQWIOVSLVFQTLQQRMDKTPLSLNTPPGEVKTLLAGCERRNAQMGCSLAGFTQ 420
|
|
|
DB 361 RRLSDNSQWIOVSLVFQTLQQRMDKTPLSLNTPPGEVKTLLAGCERRNAQMGCSLAGFTQ 420
|
|
|
QY 421 IVNEARIPACSL 432
|
|
|
DB 421 IVNEARIPACSL 432
|
|
|

```

RESULT 10

```

US-10-266-041A-1
; Sequence 1, Application US/10266041A
; Patent No. 6974690
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: PHOSPHATASES WITH IMPROVED PHYTASE ACTIVITY
; FILE REFERENCE: 19603/2798
; CURRENT APPLICATION NUMBER: US/10/266,041A
; PRIOR FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: 60/127,032
; PRIOR FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 09/540,149
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 432
; TYPE: PRF
; ORGANISM: Escherichia coli
US-10-266-041A-1

```

```

Query Match 95.1%; Score 2147; DB 2; Length 432;
Best Local Similarity 96.8%; Pred. No. 9.3e-219;
Matches 418; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

```

```

QY 1 MAAILIPFLSLIPLTPQSAFAQSEBELKLESVIVSRHGVAPPTKATQMLQDVTDPDAMP 60
|
|
|
DB 1 MAAILIPFLSLIPLTPQSAFAQSEBELKLESVIVSRHGVAPPTKATQMLQDVTDPDAMP 60
|
|
|
QY 61 TWPVGLGELTPRGSELIAVLGHYWRQLVADGLPKCCGPOSGQVAIADVDERTRKTGE 120
|
|
|
DB 61 TWPVGLGELTPRGSELIAVLGHYWRQLVADGLPKCCGPOSGQVAIADVDERTRKTGE 120
|
|
|
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOLDNANVTDAIILERAGGSIAIDFTGH 180
|
|
|
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOLDNANVTDAIILSRAGGSIAIDFTGH 180
|
|
|
QY 181 YOTAFRELERVINFPOSNTCLKREKODESCSLTQALPSELKVSADCVSLTGAIVSLASMLT 240
|
|
|
DB 181 ROTAFRELERVINFQNLCLKREKODESCSLTQALPSELKVSADNVSLTGAIVSLASMLT 240
|
|
|
QY 241 EYFLLOAOAGMPBPGRITDTSQWNTLLSLHNAQFDLLQRTPEVARSRATPLDLIKTA 300
|
|
|
DB 241 EYFLLOAOAGMPBPGRITDTSQWNTLLSLHNAQFDLLQRTPEVARSRATPLDLIKTA 300
|
|
|
QY 301 LTPHPPOKQAYGVTLPTSVLFIAGHDNTLANIGALIELNMTLPGQPDNTPPGSELVFERW 360
|
|
|
DB 301 LTPHPPOKQAYGVTLPTSVLFIAGHDNTLANIGALIELNMTLPGQPDNTPPGSELVFERW 360
|
|
|
QY 361 RRLSDNSQWIOVSLVFQTLQQRMDKTPLSLNTPPGEVKTLLAGCERRNAQMGCSLAGFTQ 420
|
|
|
DB 361 RRLSDNSQWIOVSLVFQTLQQRMDKTPLSLNTPPGEVKTLLAGCERRNAQMGCSLAGFTQ 420
|
|
|
QY 421 IVNEARIPACSL 432
|
|
|
DB 421 IVNEARIPACSL 432
|
|
|

```

RESULT 11

```

US-09-540-149A-1
; Sequence 1, Application US/09540149A
; Patent No. 651699
; GENERAL INFORMATION:

```

APPLICANT: Lei, Xingen
TITLE OF INVENTION: ENZYMES WITH IMPROVED PHYTASE ACTIVITY
FILE REFERENCE: 19603/2791
CURRENT APPLICATION NUMBER: US/09/540,149A
CURRENT FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 60/127,032
PRIOR FILING DATE: 1999-03-31
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 433
TYPE: PRT
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: UNSURE
LOCATION: (433)
OTHER INFORMATION: Xaa at position 433 in this sequence is unknown
US-09-540-149A-1

Query Match 95.1%; Score 2147; DB 2; Length 433;
Best Local Similarity 96.8%; Pred. No. 9.3e-219; Indels 0; Gaps 0;
Matches 418; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 MKAILIPFLSLILPLTPQSAFAOSEPELKESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
DB 1 MKAILIPFLSLILPLTPQSAFAOSEPELKESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
QY 61 TWPVKLGELTPRGSELIAVIGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDETRKTGE 120
DB 61 TWPVKLGELTPRGSELIAVIGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDETRKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDLPFNPLKTVGCQLDNANVTDAIIERAGGSIADEPTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDLPFNPLKTVGCQLDNANVTDAIIERAGGSIADEPTGH 180
QY 181 YQTAFRELERVLNPPQSNCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
DB 181 YQTAFRELERVLNPPQSNCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
QY 241 EIFLLQQAQMGPEPGWGRITDSHQWNTLLSLHNAQFLLQRTPEVASRATPLDLIKTA 300
DB 241 EIFLLQQAQMGPEPGWGRITDSHQWNTLLSLHNAQFLLQRTPEVASRATPLDLIKTA 300
QY 301 LTPHPQKQAVGVTLPFISVLFIAHDNTLANLGALBLNMTLPQDPNTTPGGGLVFERW 360
DB 301 LTPHPQKQAVGVTLPFISVLFIAHDNTLANLGALBLNMTLPQDPNTTPGGGLVFERW 360
QY 361 RRLSDNSQMIQVSLVFTLQOMRDKTPLSLNTPGEVKLTLAGCEERNAQMGCSLAGFTQ 420
DB 361 RRLSDNSQMIQVSLVFTLQOMRDKTPLSLNTPGEVKLTLAGCEERNAQMGCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 12
US-08-910-798-2

Sequence 2, Application US/08910798
Patent No. 5876997
GENERAL INFORMATION:
APPLICANT: KRETEZ
TITLE OF INVENTION: NOVEL PHYTASE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/910,798
FILING DATE: August 13, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: HALL, PH.D., LISA A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 09010/029001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-910-798-2

Query Match 91.4%; Score 2063.5; DB 1; Length 423;
Best Local Similarity 93.8%; Pred. No. 6.5e-210;
Matches 405; Conservative 0; Mismatches 10; Indels 17; Gaps 1;

QY 1 MKAILIPFLSLILPLTPQSAFAOSEPELKESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
DB 1 MKAILIPFLSLILPLTPQSAFAOSEPELKESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
QY 61 TWPVKLGELTPRGSELIAVIGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDETRKTGE 120
DB 61 TWPVKLGELTPRGSELIAVIGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDETRKTGE 119
QY 121 AFAAGLAPDCAITVHTQADTSSPDLPFNPLKTVGCQLDNANVTDAIIERAGGSIADEPTGH 180
DB 120 -----QADTSSPDLPFNPLKTVGCQLDNANVTDAIIERAGGSIADEPTGH 163
QY 181 YQTAFRELERVLNPPQSNCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
DB 164 YQTAFRELERVLNPPQSNCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 223
QY 241 EIFLLQQAQMGPEPGWGRITDSHQWNTLLSLHNAQFLLQRTPEVASRATPLDLIKTA 300
DB 224 EIFLLQQAQMGPEPGWGRITDSHQWNTLLSLHNAQFLLQRTPEVASRATPLDLIKTA 283
QY 301 LTPHPQKQAVGVTLPFISVLFIAHDNTLANLGALBLNMTLPQDPNTTPGGGLVFERW 360
DB 284 LTPHPQKQAVGVTLPFISVLFIAHDNTLANLGALBLNMTLPQDPNTTPGGGLVFERW 343
QY 361 RRLSDNSQMIQVSLVFTLQOMRDKTPLSLNTPGEVKLTLAGCEERNAQMGCSLAGFTQ 420
DB 344 RRLSDNSQMIQVSLVFTLQOMRDKTPLSLNTPGEVKLTLAGCEERNAQMGCSLAGFTQ 403
QY 421 IVNEARIPACSL 432
DB 404 IVNEARIPACSL 415

RESULT 13
US-09-489-039A-7512

Sequence 7512, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709,2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342

```

; SEQ ID NO 7512
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7512

```

Query Match 24.0%; Score 543; DB 2; Length 421;

Best Local Similarity 32.6%; Pred. No. 1.5e-48;

Matches 136; Conservative 70; Mismatches 181; Indels 30; Gaps 10;

```

QY 19 SAFAOSE-----PE-LKLESVIVSRHGVAP-TKATQMQDVTTPDAMPVPLGELTPR 72
DB 23 SAGAQODKAAPEGVLQGVLLMSRHNLAPLANNSSVLEGSTAKAMPQMDVPGQLTTK 82
QY 73 GGEILAYLGHYRQRLVADGLPKCCGPGSGQVAILADVDERFRKTGEAPAGLADCAI 132
DB 83 GGVLEVYWHYMEWLAQQLVYSGECPENAVYAVANSIQRTVAATQAFITGAPGCGCI 142
QY 133 TVHTQADTSSPDPLFNPPLKTGVQQLDNANVTDAI-LEBAGGSIAADPTGHYQTAPELEERV 191
DB 143 PVHHQPMQMTDPTFVITTDSPAREKALQMEKERQGMQLTE-----SYKLEETM 195
QY 192 INFPQSNCLCKREKODESCSLTQALPSBLKVSADCVSLTGAVSLASMLTEIFLLQQAQGM 251
DB 196 IDYRNSPSC-----KEKKVCSLSEKQDTFAGYQGEFVSGPLKVGSLVDAFTLOYYESF 251
QY 252 P--BEGWGRITSHQNTLISLHNAQFDLLQRTPEVARSRAFPDLITLITLTPHPQKQ 309
DB 252 PDDOVAWGEIADKQRRVLSKLNQYQDSELTSSVANAQVAKPLVYCYINALVGEASGSK- 310
QY 310 AVGVTLPTSVLFIAGHDNLANLGGALBLN-WTLPGQPDNTPGGEIVERRRLSDNQ 368
DB 311 -----AKVTLVGHDSNIALSLRLDLDPYQLPQGYRTTPIGGLLTQRMHDSAGND 363
QY 369 MIQVSLVFQTLQOMEDKTPLSINTPEGEVYKTLIAGCEERNAQMGSLAGFTQIVNEA 425
DB 364 LMKIYVYVQSTEQLNADALTLQAPQRTLALNGPVL-DDGFCPLTEFKVIVINEA 419

```

RESULT 14

US-09-489-039A-13501

; Sequence 13501, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; PRIOR FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 13501

; LENGTH: 522

; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-13501

Query Match 16.9%; Score 381; DB 2; Length 522;

Best Local Similarity 26.3%; Pred. No. 3.3e-31;

Matches 117; Conservative 73; Mismatches 209; Indels 46; Gaps 15;

```

QY 2 KAILIPLSLILPTP-GSAFAOSEPELKLSESVIVSRHGVAPTKAT-QLMQDVTTPDAW 59
DB 107 QGILRLFLTACALPLALQSAAL---ADMQLKRVLSRHGIRPTAGNREAIATGPRW 163
QY 60 PTWPVKGEELTPRGSELAYLGHYRQRLVADGLPKCCGPGSGQVAILADVDERFRKTG 119
DB 164 TEWTHDSELTHGHAAVVNNKGRABGQHRYQLGLL-QAGCPTAESIYVRAASPLQRTATYA 222
QY 120 EAFAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVQQLDNANVTDAILEBAGGSIAADFTG 179

```

```

DB 223 QALVDAFPGCGVALIHY--SGDADPLPOTDKFATQTPDPAQAAVKEAG-----DLAQ 276
QY 180 HYQTAPELEERVNLNPPQSNCLCKREKODESCSLTQALPSBLKVSADCVSLTGAVSLASML 239
DB 277 RRQA-----LAPTIQLKQAVC-----QADKPCPIFDT-PWQVQSGSKGKTTISGLSVAMN 327
QY 240 TEIFLLQQAQGM--EPGGRITDSHQNTLISLHNAQFDLLQRTPEVARSRAFPDLITDI 297
DB 328 VETLRIGSENPLSQLAMGKITQARQITALLPLLTENDVLYTAQKRGSVLHNM 387
QY 298 KTALETPHPQKQAYGVTLPVSLFIAGHDNLANLGGALBLNMTLPG-QPDNTPGGEIV 356
DB 388 LDGVPEANPNRW-----LLVADHTNLAMRTLMNPSQQLPGYSRGNIPFGSSLV 439
QY 357 FERWRRLSDNSQMIQVSLVFQTLQOMRDKTPLSINTPEGEVLT-----LACCEERNAQ 411
DB 440 LERWENASGERYRLHYFOAQGLDLDLR-----RLQTPDAQHMLQEWHPGCRQTDVGT 494
QY 412 WC-----SLAGFTQIYNEARIPACSL 432
DB 495 LCPFOALITLALGQRIIDRSSAPAVAM 519

```

RESULT 15

US-09-044-718-3

; Sequence 3, Application US/09044718

; Patent No. 6391605

; GENERAL INFORMATION:

; APPLICANT: KOSTREMA, Dirk

; APPLICANT: PASAMONTES, Luis

; APPLICANT: TONSCHY, Andrea

; APPLICANT: van LOON, Adolphus

; APPLICANT: VOGEL, Kurt

; APPLICANT: WYSS, Markus

; TITLE OF INVENTION: MODIFIED PHYTASES

; FILE REFERENCE: Modified Phytases

; CURRENT APPLICATION NUMBER: US/09/044,718

; PRIOR FILING DATE: 1998-03-19

; PRIOR APPLICATION NUMBER: EP 97810175.6

; NUMBER OF SEQ ID NOS: 82

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 3

; LENGTH: 439

; TYPE: PRT

; ORGANISM: Aspergillus fumigatus

US-09-044-718-3

Query Match 5.6%; Score 126; DB 2; Length 439;

Best Local Similarity 20.9%; Pred. No. 0.00028;

Matches 102; Conservative 69; Mismatches 166; Indels 150; Gaps 26;

```

QY 7 PFLSLILPTPOSAAOSEPELKLSESVIVSRHGVAPT-----KATQLMQDVTTPDA--- 58
DB 26 PPSLSEDELVSLSKLPK---DQRTLVVLSRHGARIVTSKSKYKKQLVATLQANAYDF 82
QY 59 -----WPTWPKVG--ELTPRGSELAYLIG--HYMRQRLVADGLPKCCGPGSGQVAIL 108
DB 83 GKPAFLKTYVYTLTGADDLTPFGSQLVNSGIKFYQRYKALARSVP-----FIR 132
QY 109 ADVDERFRKTGEAPAGI-----APDCAITVHTQADTSSPDPLFNPPLKTGV 154
DB 133 ASGSDRVIASGEKTEGQAKLADPGATNRAAPAIIVII--PESETFN-----NTLDHGV 186
QY 155 C-QLDNANVTDAILEBAGGSIAADFTGHYQTAPELEERVNLNPPQSNCLCKREKODESCSLT 213
DB 187 CTKEBASGLDGBV-----AANFTALPAPDIR-----ARAKH----- 218
QY 214 QALPSBLKVSADCVSLTGAVSLASMLTEIFLLQQAQGMPEGWRITDSHQNTLISL-- 271
DB 219 --LKGVTLDDBDVVSLMDMCSFDTV-----ARTSDASQLSPCQLFT 258
QY 272 HN--AQFDLQR-----TPE-VARSRAFPDLITLT--ALTHP 305

```

Db	259	HNEMKKNVYLQSLGKYGYGAGNPLGPAQIGFTNTELAARLTRSPYQDHTSTNSTLYSNP	318
Qy	306	PQQAIVGVTLPTSVLFLTA--GHDITNLANLGALBELAMTLPQOPDNTPPGGELVFERWRRL	363
Db	319	-----ATFPPLNATMYVDFSHDNSMWSIFPAL-----GLYNGTEPLSRTSVESAKEL	364
Qy	364	SD-NSOWI---QVSLVPOTLQOMRDKTPLSLNTPPGEVKLTLLAGCEERNAGMCSLAGFT	419
Db	365	DGYGSMWVPPGARAYFETMOCKSEKPL-VRALINDRVPLHGCDV-DKLGRCXANDFV	422
Qy	420	QIVNEAR	426
Db	423	KGLSWAR	429

Search completed: June 13, 2006, 10:30:46
 Job time : 36.6697 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2006, 10:46:00 ; Search time 9.90826 Seconds
(without alignments)
551.244 Million cell updates/sec

Title: US-10-601-319-10
Perfect score: 2258
Sequence: 1 MKAILIPFLSLILPLTPQSA.....CSLAGFTQIVNARIPACSL 432

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 64916 seqs, 12643201 residues
Total number of hits satisfying chosen parameters: 64916

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New:
1: /EMC_Celestra_SIDS3/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
2: /EMC_Celestra_SIDS3/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /EMC_Celestra_SIDS3/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /EMC_Celestra_SIDS3/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
5: /EMC_Celestra_SIDS3/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
6: /EMC_Celestra_SIDS3/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /EMC_Celestra_SIDS3/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /EMC_Celestra_SIDS3/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101.5	4.5	635	US-10-511-937-2424	Sequence 2424, Ap
2	99	4.4	384	US-11-293-697-2758	Sequence 2758, Ap
3	93.5	4.1	765	US-11-317-329-3	Sequence 3, Appl
4	93.5	4.1	765	US-11-317-329-34	Sequence 34, Appl
5	93.5	4.1	765	US-11-317-329-35	Sequence 35, Appl
6	93.5	4.1	765	US-11-317-329-36	Sequence 36, Appl
7	93.5	4.1	765	US-11-317-329-37	Sequence 37, Appl
8	93.5	4.1	765	US-11-317-329-38	Sequence 38, Appl
9	93.5	4.1	765	US-11-317-329-39	Sequence 39, Appl
10	93.5	4.1	765	US-11-317-329-40	Sequence 40, Appl
11	93.5	4.1	765	US-11-317-329-41	Sequence 41, Appl
12	93.5	4.1	765	US-11-317-329-42	Sequence 42, Appl
13	93.5	4.1	765	US-11-317-329-43	Sequence 43, Appl
14	93.5	4.1	765	US-11-317-329-44	Sequence 44, Appl
15	90.5	4.0	5738	US-10-505-928-10	Sequence 150, App
16	85.5	3.8	1127	US-10-527-411-40	Sequence 50, Appl
17	85.5	3.8	1127	US-10-527-411-50	Sequence 54, Appl
18	85.5	3.8	1127	US-10-527-411-54	Sequence 58, Appl
19	85.5	3.8	1129	US-10-527-411-58	Sequence 52, Appl
20	85.5	3.8	1129	US-10-527-411-42	Sequence 48, Appl
21	85.5	3.8	1129	US-10-527-411-48	Sequence 56, Appl
22	85.5	3.8	1129	US-10-527-411-52	Sequence 44, Appl
23	85.5	3.8	1130	US-10-527-411-56	Sequence 139, App
24	85.5	3.8	1130	US-10-527-411-44	
25	85.5	3.8	1130	US-10-527-411-139	

26	85.5	3.8	1132	6	US-10-527-411-46	Sequence 46, Appl
27	85.5	3.8	4391	7	US-11-183-325-56	Sequence 56, Appl
28	85	3.8	457	6	US-10-953-349-11735	Sequence 11735, A
29	85	3.8	460	6	US-10-953-349-34904	Sequence 34904, A
30	85	3.8	464	6	US-10-953-349-34903	Sequence 34903, A
31	85	3.8	472	6	US-10-953-349-34902	Sequence 34902, A
32	85	3.8	485	6	US-10-953-349-32471	Sequence 32471, A
33	85	3.8	504	6	US-10-953-349-32470	Sequence 32470, A
34	85	3.8	528	6	US-10-953-349-32469	Sequence 32469, A
35	85	3.8	537	6	US-10-953-349-11734	Sequence 11734, A
36	85	3.8	545	6	US-10-953-349-11733	Sequence 11733, A
37	84	3.7	847	6	US-10-505-928-300	Sequence 300, App
38	83	3.7	1730	7	US-11-270-040-6	Sequence 6, Appl
39	82	3.6	620	6	US-10-505-928-284	Sequence 284, App
40	81.5	3.6	725	7	US-11-293-697-3258	Sequence 3258, Ap
41	80.5	3.6	351	6	US-10-953-349-22620	Sequence 22620, A
42	80	3.5	760	6	US-10-505-928-50	Sequence 50, Appl
43	80	3.5	9535	6	US-10-471-571A-4496	Sequence 4496, Ap
44	79.5	3.5	800	6	US-10-953-349-33871	Sequence 33871, A
45	79.5	3.5	809	6	US-10-953-349-33870	Sequence 33870, A

ALIGNMENTS

RESULT 1
US-10-511-937-2424
Sequence 2424, Application US/10511937
Publication No. US20060088836A1
GENERAL INFORMATION:
APPLICANT: EXPRESSION DIAGNOSTICS, INC.
APPLICANT: Wohlgemuth, Jay
APPLICANT: Fry, Kirk
APPLICANT: Woodward, Robert
APPLICANT: Ly, Ngoc
APPLICANT: Prentice, James
APPLICANT: Morris, Macdonald
APPLICANT: Rosenberg, Steven
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
FILE REFERENCE: 506612000104
CURRENT APPLICATION NUMBER: US/10/511,937
CURRENT FILING DATE: 2004-10-19
PRIOR APPLICATION NUMBER: PCT/US2003/012946
PRIOR FILING DATE: 2003-04-24
PRIOR APPLICATION NUMBER: US 10/131,831
PRIOR FILING DATE: 2002-04-24
PRIOR APPLICATION NUMBER: US 10/325,899
PRIOR FILING DATE: 2002-12-20
NUMBER OF SEQ ID NOS: 3117
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2424
LENGTH: 635
TYPE: PRT
ORGANISM: Homo sapiens
US-10-511-937-2424

Query Match 4.5%; Score 101.5; DB 6; Length 635;
Best Local Similarity 20.4%; Pred. No. 0.23;
Matches 115; Conservative 60; Mismatches 195; Indels 193; Gaps 27;

QY 2 KAILIPFLSLILPLTPQSAFPAQSEF-EUKL--ESVIVS-----RHGVRAPFKAT-- 48
DB 123 RVLFPVDSGLPAPPSIIRKMGSSQPGELQISWSEBPAPISDFLRYELRYGPPDPNSTGP 182
QY 49 ---QLMQ-----DVTDPAMPVTPVKLG-----ELTPRGCE-L 76
DB 183 TVIQLIAETTCRALQRPASALDQSCAPPTPWPQDGPRTQTSBRSASALTAEQSGCL 242
QY 77 IAYL--GHWYRQ-RLVADGILLPKCGCPQSGQVAIIVADVDRTRKTRGAFAPAGLAPDCAI 132
DB 243 ISGLQPGNSYWLQURSEPDG-ISLGGSWGSGWSLPTVYDL-----PGDAVALGL----- 289

```
OY 133 TVHTQADTSSDPDLFNLKCTGVCOLDNANVTDALILERAGSIAIDFTGHYOTAFRELERVL 192
Db 290 ----QCFI-----LDLKNVTCQMQQDHASQGFTHSRAPCCPRD-- 327
OY 193 NPPOSNTCLKREK-----QDESCSLTOALPSELKVSADCVSLTGAIVSLASMLTBIFL 244
Db 328 -YPIWENCEREKNTVPLQOTPOFSRCHFKSRNDSIIHLVETVTAFTVH--SYLGSFWM 384
OY 245 LQQAQMPERG--WGRITDSH-----QNTTLLSLHNA 274
Db 385 IHQAVRLPTPLHMEISSLGHELEWQHPSSMAQETCYOLRYGEGHQDWKVLPEPLGA 444
OY 275 QDFDLQRTPEVARSRAATPLDLITKALTTPHP-----QKQAVGTLTPTSVLFI 322
Db 445 RGTLELRP---RSRYRQLPARLNGPTYGQPSWSKSDPTVERATETAMISIVTAAHLV 501
OY 323 AGHDNTLANIGALELNTLPQO-----PD-----NTPPGSELV 356
Db 502 LG-----LSAVIGLLILRWQPPAHYRLRHALWPSLPDLHRYLGQYLRDTAALSPKA--- 554
OY 357 FERRRLSDNSQMTQVSLVPTLOQMRDXTPLSLNTPPGEV-----KLTLAGCE 405
Db 555 ----TVDTCBEVEPSLL-BLPEKSSERTPLPLCSQAQMDYRRLQPSCLGTWPLSVCP 608
OY 406 ERNAGMCSLAGFTQIVNEARIP 428
Db 609 PMASGSCCT---THIANHSYLP 628
```

```
RESULT 2
US-11-293-697-2758
; Sequence 2758, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; PRIOR FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2758
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-2758
```

Query Match 4.4%; Score 99; DB 7; Length 384;

Best Local Similarity 23.0%; Pred. No. 0.18; Matches 65; Conservative 21; Mismatches 90; Indels 106; Gaps 13;

```
OY 5 LIPPLSLIIL---PLTPQSAFAQSEPELKESVIVSRHGVAPPTKATQMLQDVTTPDAMP 60
Db 69 LLPPLPTQISSSRPTPRCKSSSRPHL-----ARGRSRPLKAAQ--QDRGLPAMG 117
OY 61 TWPVTLGSLTRG--GELIATYG-----HYWRQLVADGGLPKCCCPQSG 103
Db 118 VWLSRAEVRPSALGSSRRGIDGPPSRRAAEFTQARPTKRTTPGGGRPEAGC----- 173
OY 104 QVAITADVDERTKTGBAFAAGLAPDCAITVHTQADTSSDPDLFNLKCTGVCOLDNAN-- 161
Db 174 ---WLENRQRKRRKG--ABRAGAGP---TLTPPLPAGSPDP--SPAATAAPQEPQRHL 224
OY 162 -----VTDALIERAGSIAIDFTGHYOTAFRELERVLNFPQSNLCLKREKDESCSLTQ 214
Db 225 SRGWRLLPRRRPERAASAG-----SCSRR 251
OY 215 ALPSELKVSADCVSLTGAIVSLASMLTEIFLLQQAQMPERG 256
Db 252 ALPA-----SGGAAMTGSWA-----APRPAPW 274
```

```
RESULT 3
US-11-317-329-3
; Sequence 3, Application US/11317329
; Publication No. US20060105413A1
; GENERAL INFORMATION:
; APPLICANT: Inohara, Naohiro
; TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
; FILE REFERENCE: UM-06967
; CURRENT APPLICATION NUMBER: US/11/317,329
; PRIOR FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US/10/128,174
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 765
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-317-329-3
```

Query Match 4.1%; Score 93.5; DB 7; Length 765;

Best Local Similarity 21.6%; Pred. No. 1.6; Matches 88; Conservative 46; Mismatches 144; Indels 129; Gaps 18;

```
OY 8 PLSILIPLTQSAFAQSE---PELKESVIVSRHGVAPPTKA---TOLMQDVTTPDAMP 60
Db 278 FLDTITETDILSLQSRVAVPESKALARKVSCSLSRQPREVNEDISQELMDSQNYL 337
OY 61 TWPVKGE---LTPRGSELAVLGHYW--RQRLVADGGLPKCCCPQSGQVAII---ADVD 112
Db 338 KRAQLSDRKNLVPRDELCYENKVTPLQPLVAQSGVE-----QVRLLAHVEDVD 389
OY 113 ERTKRTGBAFAAGAP-----DCAITVHTQADTSSPD-----PLFNLKCTGV 156
Db 390 CQT-----ASGYTPPLIAAQDQDPDLCALLAHGADARKVDEGWAPLHFAAQNDPDR 442
OY 157 -----LDNANVTDALIERAGSIAIDFTGHYOTAFRELERVLNFPQSNLCLKREKDESS 211
Db 443 TAILLDHACVD--QEBEGTPLHAA--ONNFEVARKLVSRQDPNLE----- 491
OY 212 LTOALPSELKVSADCVSLTGAIVSLASMLTEIFLLQQAQMPERGPGWGITDSHQNLTLSL 271
Db 492 -----ABKTPHVAAVFGHVSIVKLLT-----SGAELDAQGNLRTPL 531
OY 272 HNAQFDLQRTPEVARSRAATPLDLITKALTTPHPQKQAVG----- 312
Db 532 HLA-----VERGVRAIOHLKSGAVPDALDQSGYPLHTAARGKYLICKMLLR 581
OY 313 ---VTLPT---SVLFIAGHDNL-----ANIGALBLNMT 341
Db 582 YGASLELPTHQGWTPLHIAAYKGHLEIITHLAESHANMAGLGAIVMT 628
```

RESULT 4

US-11-317-329-34

; Sequence 34, Application US/11317329

; Publication No. US20060105413A1

; GENERAL INFORMATION:

; APPLICANT: Nunez, Gabriel

; TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling

; FILE REFERENCE: UM-06967

; CURRENT APPLICATION NUMBER: US/11/317,329

; PRIOR FILING DATE: 2005-12-22

; PRIOR APPLICATION NUMBER: US/10/128,174

; NUMBER OF SEQ ID NOS: 44

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 34

; LENGTH: 765

; TYPE: PRT

```
/ ORGANISM: Homo sapiens
US-11-317-329-34

Query Match      4.1%; Score 93.5; DB 7; Length 765;
Best Local Similarity 21.6%; Pred. No. 1.6;
Matches 88; Conservative 46; Mismatches 144; Indels 129; Gaps 18;

QY      8 FLSLIPLTPGSAFQSE---PELKESVIVISRHGVAPTKA-----TQMDQVTPDAMP 60
DB      278 FLDITETDILSLQSRVAVESKALARKVCSKSLRQPREVNEDISQELMDSGNYL 337
QY      61 TWPVVLGE---LTPRGELIAYLGHW--RQRLVADGLPKGCGPQSGOVAIT---ADVD 112
DB      338 KRALQDSRKULVPRDEBELCIYENKVTPLQFLVAGSVE-----QVRLLAHEVDVD 389
QY      113 ERTKRTGEAFAGLAP-----DCAITVHTQADTSSPD-----PLFNPLKTGVQC 156
DB      390 CQT-----ASGYTPLLIAAQDQDPDLCALLLAHGADANRVDEBGMVPLHFAAQNGDDR 442

QY      157 -----LDNANVTDALEBAGSIADFTGHYQTAFRELEVLNFPQSNLCIKREKODESCS 211
DB      443 TARLLIDHACVDA--QEREGWTPHLAA--QNNFENVARLLVSRQADPRLR-----491
QY      212 LTQALPSELKVSADCVSLTGAVSLASMLTEIFLQQAQGMPEPGWGRTDTSQWNTLSL 271
DB      492 -----ABGKTPLHVAAYFGHVSIVKLT-----SGGAEIDAAQQRNLRTPL 531
QY      272 HNAQFDLQRTPEVARSRATPLLDLTKTALTTPHPQKAYG-----312
DB      532 HLA-----VERGKVRALQHLKSGAVPDALDQSGYGPHTPAARGKYLICKMLLR 581
QY      313 -----VTLPT-----SVLFIAGHDTNL-----ANLGGALELWNT 341
DB      582 YGASLELPTHQWTPHLAAVKGHLEIITHLAESHANNGALGAVWNT 628

RESULT 5
US-11-317-329-35
/ Sequence 35, Application US/11317329
/ Publication No. US20060105413A1
/ GENERAL INFORMATION:
/ APPLICANT: Nunez, Gabriel
/ APPLICANT: Inohara, Naohiro
/ TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
/ FILE REFERENCE: UM-06967
/ CURRENT APPLICATION NUMBER: US/11/317,329
/ CURRENT FILING DATE: 2005-12-22
/ PRIOR APPLICATION NUMBER: US/10/128,174
/ PRIOR FILING DATE: 2002-04-23
/ NUMBER OF SEQ ID NOS: 44
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 35
/ LENGTH: 765
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-11-317-329-35

Query Match      4.1%; Score 93.5; DB 7; Length 765;
Best Local Similarity 21.6%; Pred. No. 1.6;
Matches 88; Conservative 46; Mismatches 144; Indels 129; Gaps 18;

QY      8 FLSLIPLTPGSAFQSE---PELKESVIVISRHGVAPTKA-----TQMDQVTPDAMP 60
DB      278 FLDITETDILSLQSRVAVESKALARKVCSKSLRQPREVNEDISQELMDSGNYL 337
QY      61 TWPVVLGE---LTPRGELIAYLGHW--RQRLVADGLPKGCGPQSGOVAIT---ADVD 112
DB      338 KRALQDSRKULVPRDEBELCIYENKVTPLQFLVAGSVE-----QVRLLAHEVDVD 389
QY      113 ERTKRTGEAFAGLAP-----DCAITVHTQADTSSPD-----PLFNPLKTGVQC 156
DB      390 CQT-----ASGYTPLLIAAQDQDPDLCALLLAHGADANRVDEBGMVPLHFAAQNGDDR 442
```

```
QY      157 -----LDNANVTDALEBAGSIADFTGHYQTAFRELEVLNFPQSNLCIKREKODESCS 211
DB      443 TARLLIDHACVDA--QEREGWTPHLAA--QNNFENVARLLVSRQADPRLR-----491
QY      212 LTQALPSELKVSADCVSLTGAVSLASMLTEIFLQQAQGMPEPGWGRTDTSQWNTLSL 271
DB      492 -----ABGKTPLHVAAYFGHVSIVKLT-----SGGAEIDAAQQRNLRTPL 531
QY      272 HNAQFDLQRTPEVARSRATPLLDLTKTALTTPHPQKAYG-----312
DB      532 HLA-----VERGKVRALQHLKSGAVPDALDQSGYGPHTPAARGKYLICKMLLR 581
QY      313 -----VTLPT-----SVLFIAGHDTNL-----ANLGGALELWNT 341
DB      582 YGASLELPTHQWTPHLAAVKGHLEIITHLAESHANNGALGAVWNT 628

RESULT 6
US-11-317-329-36
/ Sequence 36, Application US/11317329
/ Publication No. US20060105413A1
/ GENERAL INFORMATION:
/ APPLICANT: Nunez, Gabriel
/ APPLICANT: Inohara, Naohiro
/ TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
/ FILE REFERENCE: UM-06967
/ CURRENT APPLICATION NUMBER: US/11/317,329
/ CURRENT FILING DATE: 2005-12-22
/ PRIOR APPLICATION NUMBER: US/10/128,174
/ PRIOR FILING DATE: 2002-04-23
/ NUMBER OF SEQ ID NOS: 44
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 36
/ LENGTH: 765
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-11-317-329-36

Query Match      4.1%; Score 93.5; DB 7; Length 765;
Best Local Similarity 21.6%; Pred. No. 1.6;
Matches 88; Conservative 46; Mismatches 144; Indels 129; Gaps 18;

QY      8 FLSLIPLTPGSAFQSE---PELKESVIVISRHGVAPTKA-----TQMDQVTPDAMP 60
DB      278 FLDITETDILSLQSRVAVESKALARKVCSKSLRQPREVNEDISQELMDSGNYL 337
QY      61 TWPVVLGE---LTPRGELIAYLGHW--RQRLVADGLPKGCGPQSGOVAIT---ADVD 112
DB      338 KRALQDSRKULVPRDEBELCIYENKVTPLQFLVAGSVE-----QVRLLAHEVDVD 389
QY      113 ERTKRTGEAFAGLAP-----DCAITVHTQADTSSPD-----PLFNPLKTGVQC 156
DB      390 CQT-----ASGYTPLLIAAQDQDPDLCALLLAHGADANRVDEBGMVPLHFAAQNGDDR 442

QY      157 -----LDNANVTDALEBAGSIADFTGHYQTAFRELEVLNFPQSNLCIKREKODESCS 211
DB      443 TARLLIDHACVDA--QEREGWTPHLAA--QNNFENVARLLVSRQADPRLR-----491
QY      212 LTQALPSELKVSADCVSLTGAVSLASMLTEIFLQQAQGMPEPGWGRTDTSQWNTLSL 271
DB      492 -----ABGKTPLHVAAYFGHVSIVKLT-----SGGAEIDAAQQRNLRTPL 531
QY      272 HNAQFDLQRTPEVARSRATPLLDLTKTALTTPHPQKAYG-----312
DB      532 HLA-----VERGKVRALQHLKSGAVPDALDQSGYGPHTPAARGKYLICKMLLR 581
QY      313 -----VTLPT-----SVLFIAGHDTNL-----ANLGGALELWNT 341
DB      582 YGASLELPTHQWTPHLAAVKGHLEIITHLAESHANNGALGAVWNT 628

RESULT 7
US-11-317-329-37
```

Sequence 37, Application US/11317329
Publication No. US20060105413A1
GENERAL INFORMATION:
APPLICANT: Nunez, Gabriel
APPLICANT: Inohara, Naohiro
TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
FILE REFERENCE: UM-06967
CURRENT APPLICATION NUMBER: US/11/317,329
CURRENT FILING DATE: 2005-12-22
PRIOR APPLICATION NUMBER: US/10/128,174
PRIOR FILING DATE: 2002-04-23
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn version 3.1
SEQ ID NO 37
LENGTH: 765
TYPE: PRT
ORGANISM: Homo sapiens
US-11-317-329-37

Query Match 4.1%; Score 93.5; DB 7; Length 765;
Best Local Similarity 21.6%; Pred. No. 1.6; Mismatches 144; Indels 129; Gaps 18;
Matches 88; Conservative 46;

QY 8 FLSDLLPLTPQSAFAQSE---PELKLESVIVSRHGVRAPTKA---TQLMQDYTPDAMP 60
DB 278 FLDITIEDILLSLQSRVAVPESKALARKVSCSLSRQPREVNEDIQELMDSDSGNYL 337
QY 61 TWPKVLGE---LTPRGSELIAVLGHYV-RQRLVADGILLPKGCCQSGQOVAIL---ADVD 112
DB 338 KRALQSLDRKNLVPRDEBELCIYENKVTPLQPLVAQSVB---QVRLLLAHEVDVD 389
QY 113 ERTKRTGEAFAGAP-----DCAITVHTQADTSSPD-----PLFNPPLKTGVQC 156
DB 390 CQT-----ASGYTPLLIAAQDQPDLCALLAHGADARVDEMGAPLHFAQNGDDR 442
QY 157 ---LDNANYTDALIERAGSIADFTGHYQTAFLERELVINFPQSNLCLKREKODESCS 211
DB 443 TABLLDHGACVDA-QEREGWTPPLHLAA--QNNFENVARLLVSRQADPNLRE----- 491
QY 212 LTQALPSELKVSADCVSLTGAVSLASMTETIFLLQQAQMPBPCKGRITTDHQNNTLLSL 271
DB 492 ---AEGKTPPLHVAAYFGHVSIVKLLT-----SQGAEILDQQRNLRTPL 531
QY 272 HNAQFDLLQRTPEVARSRATPLDLIKTALTPHPPOKQAYG----- 312
DB 532 HLA-----VERGKVRALIQHLKSGAVPDALDQSGYGPPLHTAARCKYLICKMLLR 581
QY 313 ---VTLPT---SVLPLAGHDNTL-----ANLGALELNMWT 341
DB 582 YGASLELPTHQGWTPPLHLAAVKGHLEIHLAESHANMGALGAVNMT 628

RESULT 8
US-11-317-329-38
Sequence 38, Application US/11317329
Publication No. US20060105413A1
GENERAL INFORMATION:
APPLICANT: Nunez, Gabriel
APPLICANT: Inohara, Naohiro
TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
FILE REFERENCE: UM-06967
CURRENT APPLICATION NUMBER: US/11/317,329
CURRENT FILING DATE: 2005-12-22
PRIOR APPLICATION NUMBER: US/10/128,174
PRIOR FILING DATE: 2002-04-23
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn version 3.1
SEQ ID NO 38
LENGTH: 765
TYPE: PRT
ORGANISM: Homo sapiens
US-11-317-329-38

Query Match 4.1%; Score 93.5; DB 7; Length 765;
Best Local Similarity 21.6%; Pred. No. 1.6; Mismatches 144; Indels 129; Gaps 18;
Matches 88; Conservative 46;

QY 8 FLSDLLPLTPQSAFAQSE---PELKLESVIVSRHGVRAPTKA---TQLMQDYTPDAMP 60
DB 278 FLDITIEDILLSLQSRVAVPESKALARKVSCSLSRQPREVNEDIQELMDSDSGNYL 337
QY 61 TWPKVLGE---LTPRGSELIAVLGHYV-RQRLVADGILLPKGCCQSGQOVAIL---ADVD 112
DB 338 KRALQSLDRKNLVPRDEBELCIYENKVTPLQPLVAQSVB---QVRLLLAHEVDVD 389
QY 113 ERTKRTGEAFAGAP-----DCAITVHTQADTSSPD-----PLFNPPLKTGVQC 156
DB 390 CQT-----ASGYTPLLIAAQDQPDLCALLAHGADARVDEMGAPLHFAQNGDDR 442
QY 157 ---LDNANYTDALIERAGSIADFTGHYQTAFLERELVINFPQSNLCLKREKODESCS 211
DB 443 TABLLDHGACVDA-QEREGWTPPLHLAA--QNNFENVARLLVSRQADPNLRE----- 491
QY 212 LTQALPSELKVSADCVSLTGAVSLASMTETIFLLQQAQMPBPCKGRITTDHQNNTLLSL 271
DB 492 ---AEGKTPPLHVAAYFGHVSIVKLLT-----SQGAEILDQQRNLRTPL 531
QY 272 HNAQFDLLQRTPEVARSRATPLDLIKTALTPHPPOKQAYG----- 312
DB 532 HLA-----VERGKVRALIQHLKSGAVPDALDQSGYGPPLHTAARCKYLICKMLLR 581
QY 313 ---VTLPT---SVLPLAGHDNTL-----ANLGALELNMWT 341
DB 582 YGASLELPTHQGWTPPLHLAAVKGHLEIHLAESHANMGALGAVNMT 628

RESULT 9
US-11-317-329-39
Sequence 39, Application US/11317329
Publication No. US20060105413A1
GENERAL INFORMATION:
APPLICANT: Nunez, Gabriel
APPLICANT: Inohara, Naohiro
TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
FILE REFERENCE: UM-06967
CURRENT APPLICATION NUMBER: US/11/317,329
CURRENT FILING DATE: 2005-12-22
PRIOR APPLICATION NUMBER: US/10/128,174
PRIOR FILING DATE: 2002-04-23
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn version 3.1
SEQ ID NO 39
LENGTH: 765
TYPE: PRT
ORGANISM: Homo sapiens
US-11-317-329-39

Query Match 4.1%; Score 93.5; DB 7; Length 765;
Best Local Similarity 21.6%; Pred. No. 1.6; Mismatches 144; Indels 129; Gaps 18;
Matches 88; Conservative 46;

QY 8 FLSDLLPLTPQSAFAQSE---PELKLESVIVSRHGVRAPTKA---TQLMQDYTPDAMP 60
DB 278 FLDITIEDILLSLQSRVAVPESKALARKVSCSLSRQPREVNEDIQELMDSDSGNYL 337
QY 61 TWPKVLGE---LTPRGSELIAVLGHYV-RQRLVADGILLPKGCCQSGQOVAIL---ADVD 112
DB 338 KRALQSLDRKNLVPRDEBELCIYENKVTPLQPLVAQSVB---QVRLLLAHEVDVD 389
QY 113 ERTKRTGEAFAGAP-----DCAITVHTQADTSSPD-----PLFNPPLKTGVQC 156
DB 390 CQT-----ASGYTPLLIAAQDQPDLCALLAHGADARVDEMGAPLHFAQNGDDR 442
QY 157 ---LDNANYTDALIERAGSIADFTGHYQTAFLERELVINFPQSNLCLKREKODESCS 211
DB 443 TABLLDHGACVDA-QEREGWTPPLHLAA--QNNFENVARLLVSRQADPNLRE----- 491

```
QY 212 LTQALPSBLKVSADCVSLTGAIVSLASMTLEIFLLQQAQGMPEKGRITDSHQWTLTSL 271
DB 442 -----ABEKTPLHVAAYFGHVSIVKLT-----SQGAELDAQORNLRTPL 531
QY 272 HNAQFDLQRTPEVARSBATPLDLIKTALTTPHPQKQAYG----- 312
DB 532 HLA-----VERGKVRALQHLKSGAVPDALDQSGYGPLHTAARGKYLICKMLLR 581
QY 313 -----VTLPF-----SVLFIAGHDTNL-----ANIGALELWMT 341
DB 582 YGASLELPTHQWMTPLHLAAVKGHLEIHLAESHANNMGAALGAVWMT 628

RESULT 10
US-11-317-329-40
; Sequence 40, Application US/11317329
; Publication No. US20060105413A1
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
; FILE REFERENCE: UM-06967
; CURRENT APPLICATION NUMBER: US/11/317,329
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US/10/128,174
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40
; LENGTH: 765
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-317-329-40

Query Match 4.1%; Score 93.5; DB 7; Length 765;
Best Local Similarity 21.6%; Pred. No. 1.6;
Matches 88; Conservative 46; Mismatches 144; Indels 129; Gaps 18;

QY 8 FLSLIPLTPGSAFPAQSE---PELKLESVIVSRHGVARPTKA---TQLMODVTPDAMP 60
DB 278 FLDITERTDILSLQSRVAVPESKALARKVSCKSLRQPREVNEDISQELMDSQNYL 337
QY 61 TWPVVLGE---LTPRGELIAYLGHYW---RQRLVADGGLPKGCCPGSGQVAII---ADVD 112
DB 338 KRALQLSRKNLVPDEBELCIYENKVTPLQFLVAQGSVE-----QVRLLAHEVDVD 389
QY 113 ERTKTGEAFAGLAP-----DCATVHTQADTSSPD-----PLFNPILKTGYCQ 156
DB 390 CQT-----ASGYTPPLIAADQDPDLCALLLHAGADANRVDDEGMARPLHRAQNGDDR 442
QY 157 -----LDNANVTDAILEPAGSGIADFTGHYQTAPELEBVLNFPQSNICLKREKODESCS 211
DB 443 TARRLLDGGACVDA--QEREQWTPPLHLAA--QNNFENVARLLVSRQADPVLRE----- 491
QY 212 LTQALPSBLKVSADCVSLTGAIVSLASMTLEIFLLQQAQGMPEKGRITDSHQWTLTSL 271
DB 492 -----ABEKTPLHVAAYFGHVSIVKLT-----SQGAELDAQORNLRTPL 531
QY 272 HNAQFDLQRTPEVARSBATPLDLIKTALTTPHPQKQAYG----- 312
DB 532 HLA-----VERGKVRALQHLKSGAVPDALDQSGYGPLHTAARGKYLICKMLLR 581
QY 313 -----VTLPF-----SVLFIAGHDTNL-----ANIGALELWMT 341
DB 582 YGASLELPTHQWMTPLHLAAVKGHLEIHLAESHANNMGAALGAVWMT 628

RESULT 11
US-11-317-329-41
; Sequence 41, Application US/11317329
; Publication No. US20060105413A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
; FILE REFERENCE: UM-06967
; CURRENT APPLICATION NUMBER: US/11/317,329
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US/10/128,174
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
; LENGTH: 765
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-317-329-41

Query Match 4.1%; Score 93.5; DB 7; Length 765;
Best Local Similarity 21.6%; Pred. No. 1.6;
Matches 88; Conservative 46; Mismatches 144; Indels 129; Gaps 18;

QY 8 FLSLIPLTPGSAFPAQSE---PELKLESVIVSRHGVARPTKA---TQLMODVTPDAMP 60
DB 278 FLDITERTDILSLQSRVAVPESKALARKVSCKSLRQPREVNEDISQELMDSQNYL 337
QY 61 TWPVVLGE---LTPRGELIAYLGHYW---RQRLVADGGLPKGCCPGSGQVAII---ADVD 112
DB 338 KRALQLSRKNLVPDEBELCIYENKVTPLQFLVAQGSVE-----QVRLLAHEVDVD 389
QY 113 ERTKTGEAFAGLAP-----DCATVHTQADTSSPD-----PLFNPILKTGYCQ 156
DB 390 CQT-----ASGYTPPLIAADQDPDLCALLLHAGADANRVDDEGMARPLHRAQNGDDR 442
QY 157 -----LDNANVTDAILEPAGSGIADFTGHYQTAPELEBVLNFPQSNICLKREKODESCS 211
DB 443 TARRLLDGGACVDA--QEREQWTPPLHLAA--QNNFENVARLLVSRQADPVLRE----- 491
QY 212 LTQALPSBLKVSADCVSLTGAIVSLASMTLEIFLLQQAQGMPEKGRITDSHQWTLTSL 271
DB 492 -----ABEKTPLHVAAYFGHVSIVKLT-----SQGAELDAQORNLRTPL 531
QY 272 HNAQFDLQRTPEVARSBATPLDLIKTALTTPHPQKQAYG----- 312
DB 532 HLA-----VERGKVRALQHLKSGAVPDALDQSGYGPLHTAARGKYLICKMLLR 581
QY 313 -----VTLPF-----SVLFIAGHDTNL-----ANIGALELWMT 341
DB 582 YGASLELPTHQWMTPLHLAAVKGHLEIHLAESHANNMGAALGAVWMT 628

RESULT 12
US-11-317-329-42
; Sequence 42, Application US/11317329
; Publication No. US20060105413A1
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
; FILE REFERENCE: UM-06967
; CURRENT APPLICATION NUMBER: US/11/317,329
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US/10/128,174
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 765
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-317-329-42

Query Match 4.1%; Score 93.5; DB 7; Length 765;
Best Local Similarity 21.6%; Pred. No. 1.6;
Matches 88; Conservative 46; Mismatches 144; Indels 129; Gaps 18;
```


/ CURRENT APPLICATION NUMBER: US/10/505,928
/ CURRENT FILING DATE: 2004-08-27
/ PRIOR APPLICATION NUMBER: US 60/363,019
/ PRIOR FILING DATE: 2002-03-07
/ NUMBER OF SEQ ID NOS: 866
/ SOFTWARE: PatentIn 3.2
/ SEQ ID NO 150
/ LENGTH: 5738
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-10-505-928-150

Query Match 4.0%; Score 90.5; DB 6; Length 5738;
Best Local Similarity 22.9%; Pred. No. 53;

Matches 110; Conservative 35; Mismatches 159; Indels 177; Gaps 29;

```
QY      4 ILIPFLSLILIPLPQSAF-----AQSBEYLKASVIVSRHGR---APTATQLMOD 53
DB      2359 VVLBELQPRRLRRPRVRPMNRPGQTBEBCSGCHGQVNRVGERMNGRCRCVCCSLN 2418
QY      54 VTRDAMPTWPKLGLTRPGSELIVLGHYWRQLVADGLLRKCGCRGQVAIIVDVE 113
DB      2419 LTAHCSRYCP--LGS-CPOGMVLVEGTGR-----SCCHCALPGE----- 2454
QY      114 RTRKTGEAFAGIARPCATVNHQADTSSPDPLFN-ELKTVGQQLDNANVTDAILERAG 172
DB      2455 --NOTQPRNA-----TPAARAPSPQIRPLAT-----YILPPSG 2488
QY      173 SIADFTGHYOTAFRELERYLNPQSNLCRKREKODECSLTQALPSELKVSADCVSLTGA 232
DB      2489 SC-----RPLSSP-TPACL-----SLNHRP-----CYSPLGI 2515
QY      233 VSLA-----SMLTEIFLL-QQAQGMREPGWGRITDSH-QMNTLLSLHNAQFDLL 279
DB      2516 AGIAGSGILHASSCOLEHPTQALLGAPYQSPSPQGMHAGDAVAKMHT--RPHYLQDL 2573
QY      280 Q-----RTPVARSRA-----TPLLDLIKALTPHRPKQAYG---- 312
DB      2574 QPRNLTGILVPEBTSSNAYASSPSIQSSNGLHMDYRDL-PGILPLPKVSPAQGRWQ 2632
QY      313 -VTLFTSVLFIAGHDNTNLNLGGALEIANTLPGQPDNTPRG-----GELVFER-WRR 362
DB      2633 QPTWP-----FCGFHS-----LCFQGPSSVPEGHGHSMLVETLLFRNWD 2674
QY      363 LSDNSQWIVSLVFTQLQMR--DKTPSLNTPRGSEVKLTLAGCEB---RNAQGMCSLA 416
DB      2675 L-DPAVM-----TFGRMVQARFVRVWRHVDVHSDVPLQVELLGCSEPGVGLRCASGECVLR 2728
QY      417 G 417
DB      2729 G 2729
```

Search completed: June 13, 2006, 10:51:03
Job time : 10.9083 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2006, 10:22:24 ; Search time 25.2661 Seconds

(without alignments)
1645.118 Million cell updates/sec

Title: US-10-601-319-10

Perfect score: 2258

Sequence: 1 MKAILIPFLSLILPLTPQSA.....CSLAGFTQIVNEARIPACSL 432

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2182	96.6	432	2	B36733 acid phosphatase (
2	2144	95.0	434	2	H90770 phosphoanhydride p
3	2144	95.0	444	2	D85633 hypothetical prote
4	959.5	42.5	441	2	AC0201 acid phosphatase (
5	560	24.8	413	2	F90773 periplasmic glucos
6	554	24.5	413	2	B85636 periplasmic glucos
7	547	24.2	413	2	UV0087 glucose-1-phosphat
8	538	23.8	417	2	S25627 glucose-1-phosphat
9	536	23.7	414	2	B87316 periplasmic phosph
10	536	23.7	414	2	B87316 periplasmic phosph
11	137	6.1	416	2	T16058 hypothetical prote
12	132	5.8	423	1	S06167 acid phosphatase (
13	116.5	5.2	423	1	A33395 acid phosphatase (
14	112.5	5.0	428	2	S64682 acid phosphatase (
15	111.5	4.9	421	2	S14742 acid phosphatase (
16	109.5	4.8	344	2	B89130 protein F52B1.8 (1
17	109.5	4.8	457	2	A56925 paired box transcr
18	109.5	4.8	479	1	JN0890 acid phosphatase (
19	108	4.8	386	1	JN0610 acid phosphatase (
20	107.5	4.8	459	2	S52520 paired box transcr
21	107.5	4.8	479	1	JN0715 paired box transcr
22	107	4.7	5126	2	S40450 3-phenylase (BC 3.1.
23	105	4.7	450	2	A54429 ryanodine receptor
24	104.5	4.6	397	2	C81716 paired box transcr
25	103.5	4.6	537	2	S54770 hypothetical prote
26	103.5	4.6	888	2	T46726 secreted acid phos
27	103.5	4.6	1048	1	BVCS6C exonuclease (BC 3.
28	103	4.6	769	2	B87681 tyrosine kinase Di
29	101.5	4.5	356	2	F75594 hypothetical prote

30	101.5	4.5	635	2	A45266	MP-L-P protein prec
31	101.5	4.5	789	2	AE0064	DNA-directed DNA p
32	101.5	4.5	1214	2	AG2897	conserved hypothet
33	101.5	4.5	1387	2	A97673	probable periplasm
34	100.5	4.5	425	2	D83186	hypothetical prote
35	100	4.4	971	2	T32883	hypothetical prote
36	99.5	4.4	234	2	F42696	thrombin (BC 3.4.2
37	99	4.4	560	2	F84341	hypothetical prote
38	99	4.4	969	2	T27977	hypothetical prote
39	99	4.4	1413	2	D86844	protein ZK792.1 (1
40	98.5	4.4	381	2	JH0152	acid phosphatase (
41	98	4.3	1034	2	AB0551	exonuclease SpCC (
42	97.5	4.3	4056	2	H96599	pyridoxal phosphat
43	96.5	4.3	328	2	A83572	subtilisin (EC 3.4
44	96.5	4.3	374	2	139781	ATP-dependent dsbN
45	96.5	4.3	1047	2	C85535	

ALIGNMENTS

RESULT 1
B36733
acid phosphatase (EC 3.1.3.2) precursor - Escherichia coli (strain K-12)
N:Alternate names: phosphoanhydride phosphohydrolyase, periplasmic, phytase 6
C:Species: Escherichia coli
C:Date: 19-Apr-1991 #sequence revision 19-Apr-1991 #text change 09-Jul-2004
C:Accession: B36733; S18018; B64839; A26534; S17960; S33278
R:Daase, J.; Marck, C.; Boquet, P.L.
J. Bacteriol. 172, 5497-5500, 1990
A:Title: The complete nucleotide sequence of the Escherichia coli gene appa reveals sig.
A:Reference number: A36733; MUID:90368616; PMID:2168385
A:Accession: B36733
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-432 <DAS>
A:Cross-references: UNIPROT: P07102; UNIPARC: UP1000004750A; GB: M58708; NID: g145283; PIDN: R191.
R:Greiner, R.; Tany, K.D.
Biol. Chem. Hoppe-Seyler 372, 664-665, 1991
A:Title: Characterization of a phytase from Escherichia coli.
A:Reference number: S18018
A:Accession: S18018
A:Molecule type: protein
A:Residues: 23-33 <GRS>
A:Cross-references: UNIPARC: UP10000179815
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID: 97426617; PMID: 9278503
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-432 <BLAT>
A:Cross-references: UNIPARC: UP1000004750A; GB: AE000200; GB: U00096; NID: g2367111; PIDN: AA
R: Touati, R.; Danchin, A.
Biochimie 69, 215-221, 1987
A:Title: The structure of the promoter and amino terminal region of the pH 2.5 acid phos
A:Reference number: A26534; MUID: 87271766; PMID: 3038201
A:Accession: A26534
A:Molecule type: DNA
A:Residues: 1-50, 'NAGCHPRMANLAC', 65, 'T', 67-74, 'DV', 77-111, 'S', <TOU>
A:Cross-references: UNIPARC: UP1000016BDE2; GB: X05471; NID: g40925; PIDN: CAA29031.1; PID: g
R: Daase, J.; Fahl, H.; Marck, C.; Dion, M.; Kieffer-Bontemps, M.; Boquet, P.L.
Mol. Gen. Genet. 229, 341-352, 1991
A:Title: A new oxygen-regulated operon in Escherichia coli comprises the genes for a put
A:Reference number: S17958; MUID: 92045231; PMID: 1658595
A:Accession: S17958
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-17 <DA2>
A:Cross-references: UNIPARC: UP1000016FPB1; GB: S63811; NID: g238656; PIDN: AAB20286.1; PID:

R:Greiner, R.; Konietzny, U.; Jany, K.D.
Arch. Biochem. Biophys. 303, 107-113, 1993
A:Title: Purification and characterization of two phytaes from *Escherichia coli*.
A:Reference number: S33278; MUID:93256556; PMID:8387749
A:Accession: S33278
A:Molecule type: protein
A:Residues: 23-31, 'A', 33-34 <GR2>
A:Cross-references: UNIPARC:UPI0000179816
C:Comment: In addition to cAMP-mediated control, this enzyme is induced when bacterial c
naerobic conditions.
C:Genetics:
A:Gene: appa
A:Map position: 25 min
C:Keywords: monomer; periplasmic space; phosphohistidine; phosphoprotein; phosphoric mor
F:1-32/Domain: signal sequence #status predicted <STG>
F:23-112/Product: acid phosphatase #status predicted <MAT>
F:38/Active site: Arg #status predicted
F:39/Active site: His (phosphohistidine intermediate) #status predicted

Query Match 96.6%; Score 2182; DB 2; Length 432;
Best Local Similarity 98.1%; Pred. No. 2e-167;
Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 MKAILIPLSLILPLTPQSAFAQS-EPBLKLSVIVSRHGVAPFKATQIMODVTPDAMP 60
Db 1 MKAILIPLSLILPLTPQSAFAQS-EPBLKLSVIVSRHGVAPFKATQIMODVTPDAMP 60

Qy 61 TWPVKGLGELTPRGELIAYLGHYRQRLVADGLPKCGCGQGVAAIADVDERTKKTGE 120
Db 61 TWPVKGLGELTPRGELIAYLGHYRQRLVADGLPKCGCGQGVAAIADVDERTKKTGE 120

Qy 121 APAAGLADPCATVHTQADTSSPDLFNPILKTGVQOLNANVTDAILSRAGSIADFTGH 180
Db 121 APAAGLADPCATVHTQADTSSPDLFNPILKTGVQOLNANVTDAILSRAGSIADFTGH 180

Qy 181 YQTAARELERVLPQSNLCLREKQDSCSLTQALPSELKVSADCVSLTGAVSLASM 240
Db 181 YQTAARELERVLPQSNLCLREKQDSCSLTQALPSELKVSADCVSLTGAVSLASM 240

Qy 241 EIFLIQQAQMGPEPGMGRTTDSHQNNTLLSLHNAQFDLLQRTPEVARSRATPLDLIKTA 300
Db 241 EIFLIQQAQMGPEPGMGRTTDSHQNNTLLSLHNAQFDLLQRTPEVARSRATPLDLIKTA 300

Qy 301 LTPHPPOKQAGVTLPTSVLFTAGHDTNLANIGALELNMWLPQGPDPNTPPGSELVFERW 360
Db 301 LTPHPPOKQAGVTLPTSVLFTAGHDTNLANIGALELNMWLPQGPDPNTPPGSELVFERW 360

Qy 361 RRLSDNSQMIQVSLVFQTLQQRKDTPLSLNTPPGSEVLTLAGCERNAQMGCSLAGFTQ 420
Db 361 RRLSDNSQMIQVSLVFQTLQQRKDTPLSLNTPPGSEVLTLAGCERNAQMGCSLAGFTQ 420

Qy 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432

RESULT 2
H90770
phosphoanhydride phosphorylae [imported] - *Escherichia coli* (strain O157:H7, substrain
C:Species: *Escherichia coli*
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: H90770
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasaara, N.; Yasunaga, T.; Kuhara, S.; Shibata, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genc
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: H90770
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-334 <HNA>
A:Cross-references: UNIPROT:O8XC29; UNIPARC:UPI000000517; GB:BA000007; PIDN:BA034559.1;
A:Experimental source: strain O157:H7, substrain RMD 0509952

C:Genetics:
A:Gene: ECG136

Query Match 95.0%; Score 2144; DB 2; Length 434;
Best Local Similarity 96.5%; Pred. No. 2.3e-164;
Matches 419; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

Qy 1 MKAILIPLSLILPLTPQSAFAQS-EPBLKLSVIVSRHGVAPFKATQIMODVTPDA 58
Db 1 MKAILIPLSLILPLTPQSAFAQS-EPBLKLSVIVSRHGVAPFKATQIMODVTPDA 60

Qy 59 WPTVPVKGELTPRGELIAYLGHYRQRLVADGLPKCGCGQGVAAIADVDERTKKT 118
Db 59 WPTVPVKGELTPRGELIAYLGHYRQRLVADGLPKCGCGQGVAAIADVDERTKKT 120

Qy 119 GEAPAGLADPCATVHTQADTSSPDLFNPILKTGVQOLNANVTDAILSRAGSIADFT 178
Db 121 GEAPAGLADPCATVHTQADTSSPDLFNPILKTGVQOLNANVTDAILSRAGSIADFT 180

Qy 179 GHYQTAARELERVLPQSNLCLREKQDSCSLTQALPSELKVSADCVSLTGAVSLASM 238
Db 181 GHYQTAARELERVLPQSNLCLREKQDSCSLTQALPSELKVSADCVSLTGAVSLASM 240

Qy 239 LTESIFLIQQAQMGPEPGMGRTTDSHQNNTLLSLHNAQFDLLQRTPEVARSRATPLDLIK 298
Db 241 LTESIFLIQQAQMGPEPGMGRTTDSHQNNTLLSLHNAQFDLLQRTPEVARSRATPLDLIK 300

Qy 299 TALTPHPPOKQAGVTLPTSVLFTAGHDTNLANIGALELNMWLPQGPDPNTPPGSELVFE 358
Db 301 TALTPHPPOKQAGVTLPTSVLFTAGHDTNLANIGALELNMWLPQGPDPNTPPGSELVFE 360

Qy 359 RWRRLSDNSQMIQVSLVFQTLQQRKDTPLSLNTPPGSEVLTLAGCERNAQMGCSLAGF 418
Db 361 RWRRLSDNSQMIQVSLVFQTLQQRKDTPLSLNTPPGSEVLTLAGCERNAQMGCSLAGF 420

Qy 419 TOIVNEARIPACSL 432
Db 421 TOIVNEARIPACSL 434

RESULT 3
D85633
hypothetical protein appa [imported] - *Escherichia coli* (strain O157:H7, substrain EDL93
C:Species: *Escherichia coli*
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: D85633
R:Perma, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: D85633
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-444 <STO>
A:Cross-references: UNIPROT:O8XC29; UNIPARC:UPI0000165751; GB:AE005174; NID:g12514245; P
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: appa

Query Match 95.0%; Score 2144; DB 2; Length 444;
Best Local Similarity 96.5%; Pred. No. 2.4e-164;
Matches 419; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

Qy 1 MKAILIPLSLILPLTPQSAFAQS-EPBLKLSVIVSRHGVAPFKATQIMODVTPDA 58
Db 11 MKAILIPLSLILPLTPQSAFAQS-EPBLKLSVIVSRHGVAPFKATQIMODVTPDA 70

Qy 59 WPTVPVKGELTPRGELIAYLGHYRQRLVADGLPKCGCGQGVAAIADVDERTKKT 118
Db 71 WPTVPVKGELTPRGELIAYLGHYRQRLVADGLPKCGCGQGVAAIADVDERTKKT 130

Qy 119 GEAPAGLADPCATVHTQADTSSPDLFNPILKTGVQOLNANVTDAILSRAGSIADFT 178

Query Match	24.8%;	Score 560;	DB 2;	Length 413;
Best Local Similarity	33.3%;	Pred. No. 3.7e-317;		
Matches 141;	Conservative	73;	Mismatches 174;	Indels 36;
			Gaps	12;
QY	13	IPLEPQSAFAOSEPE-LKLESVIVSRHGVRAP-TKATQLMDVTPDAMPYVGLGELT	70	
Db	13	VVLASNNQAQVPEGYQLQOQVLAHMSRHNLRAPLANNNGSVLEQSTPNKPEMDVPGQLT	72	
QY	71	PRGSELIALYHTYMQRLVADGSLPKCCGPOSGOYAITADYDERTKTKGEAPAGLAPDC	130	
Db	73	TKGGLLEVYMGHYMKYEMWLEAQGMVKSGBCPDPDYAVANSLQRVAATAQFFITGAPGCG	132	
QY	131	AITVYTAQDTSPPDELFPNPLKTVGCQLDNANYTD---AILEPAGSGIADFGHQTARPE	187	
Db	133	DIPVHQRKMGMDTFNFPVITD---DSAAFSEBAVAAHKEKLSKL-QLDTSQVL-----	183	
QY	188	LERVINFPOSNICLRKRYQDBSCSLTQALPSELKVSADCVSLTGAVSLASMTLPIQLQ	247	
Db	184	LEKIVNYXDSPAC---KEKQCCSLVPDGKNTFSAKYQOBPGVSGPLKAGNSLVNAFTLYQY	239	
QY	248	AQGMPE-EPGMRITDSSHOMNTLSLHNAQFDLQRTPEVARSRPTPLDILKTALTPHP	305	
Db	240	YEGFPMDOVANGELISDQOMKVLSTLAKNGYQDSLFTSEVARNAYAKPLVSTIIDRL	295	
QY	306	POKQAYGVTLPTS---VLFIAHDTNLIANGALELN-WTLPGQPDNTPPGSELVFERMR	361	
Db	296	-----VTDRTSAPKRTIVLVGHDSNISIALLTALDFPKFYQLHDQNERPIGAKIVFGMR	348	
QY	362	RLSDNSQMIQVSLVEQTLQOMRDKTLP.LSLNTPRGSEVTKTLIAGCERNNAQMGCSLAGFTQI	421	
Db	349	DSKANRDLMKLEIVYQSAEQLRNADALTLQAPQAVRTLELSGC-PIADGCFPMDKFDSV	407	
QY	422	VNEA 425		
Db	408	INEA 411		

[illegible]

Query Match 24.2%; Score 547; DB 2; Length 413;
Best Local Similarity 33.7%; Pred. No. 4, 1e-36;
Matches 138; Conservative 66; Mismatches 181; Indels 24; Gaps 10;

QY 22 AQSEPE-LKLSVSVYSRHGVAP-TKATQLMQDVTPDAMPPTVPLGELTREGELIAY 79
DB 22 AQTEBEGYQLOQVLMMSRNLRLAPLANNGNVLAQSTPNAMPAMDVPGQLTKGGLLEVY 81
QY 80 LGHYRQRLVADGLLPKCCGPOGQVATADDERTRKTGEAPAGLAPDCATITHTQAD 139
DB 82 MCHYRREMLVAGGLLPGECPADVTLYAVANSIQRTVAFOFTISAPFGCOIPVHQRK 141
QY 140 TSSPPLFNPPLKTYGCOLDNANVTDAILERAGGSIAADFTGHYQTAPELERVLNFPQSNL 199
DB 142 MWTNPTFRPVITTDSSAARQAVQV-MEKASQ-----HIDESTKLEQITHTQDSFS 195
QY 200 CLKREKODESCSLTQALPSELKVSADCVSLTGAVALSMLETEIFLLOAQAGMP--EPGWG 257
DB 196 C----KEKHQCSLIDAKDTFSANYQOEPGVQPLKGNGLVDAFTLQYVEGFPMQVAMG 251
QY 258 RTDHSQWNTLLSLHNAQFDLQRTPEVARSATPLDLIKTALTPHPQKQAYGVTLFT 317
DB 252 GHTDRQWNVLSKLKNGYQDSLFTSPYARNVAAPLVKYIDKVLV-----AERVSAF- 303
QY 318 SVLFIAGHDYTNLNGALELN-WTLPGQPDNTPGEGELVFERWRRLSDNSQMIQVSLVF 376
DB 304 KVTVLVGHDSNLASLITLADFPYQLHDYERIPGGQLVFRMDGNANRDLMKIEVY 363
QY 377 QTLQMRDTPPLSLNTPPEGEVLTLAGCEERNAQGCISLAGEFTQIVNEA 425
DB 364 QSARQLRNAEALTLKSPAQRTVLELKGCPV-DANGFCPLDKDYNMNTA 411

RESULT 9

S25627
Glucose-1-phosphatase (EC 3.1.3.10) precursor - Providencia rettgeri
C.Species: Providencia rettgeri
C.Date: 25-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C.Accession: S25627
R.Riccio, M.L.; Chiesurin, A.; Lombardi, G.; Satta, G.
submitted to the EMBL Data Library, September 1992
A.Reference number: S25627
A.Accession: S25627
A.Molecule type: DNA
A.Residues: 1-417 <RIC>
A.Cross-references: UNIPROT:Q52309; UNIPARC:UPI00001256FC; EMBL:X68201; NID:G45771; PIDN
A.Experimental source: strain PV7
C.Genetics:
A.Gene: agp
C.Keywords: periplasmic space; phosphohistidine; phosphoprotein; phosphoric monoester hy
F.1-39/Domain: signal sequence #status predicted <Sig>
F.30-417/Product: glucose-1-phosphatase #status predicted <Mat>
F.42/Active site: His (phosphohistidine intermediate) #status predicted
F.314/Active site: His #status predicted

Query Match 23.8%; Score 538; DB 2; Length 417;
Best Local Similarity 30.4%; Pred. No. 2, 2e-35;
Matches 130; Conservative 85; Mismatches 175; Indels 38; Gaps 11;

QY 10 SLILPTQSAFAQSEPELKLSVSVYSRHGVAPTKATQLMQDVTPDAMPPTVPLGEL 69
DB 13 ALFAPAIATPMASSTNDQAMVDQVAVLSRHNLRTIYVNTGILTEVTDKMPMDAKSGYL 72
QY 70 TPRGSELIAVIGHYWRQ-----RLVADGLLPKCCGPOGQV-VATADDERTRKTGEAPA 123
DB 73 ITQCGALVYMGHYFRWIDQNKLADEL-----CPTSNEDIYLYTNSLQRTIATAQFA 127
QY 124 AGLAADCAITVHTQADTSSPDLFNPPLKTYGCOLDNANVTDAILERAGGSIAADFTG--HY 181
DB 128 AGAFGCKVNIHQEIGMDVFNPIIT-----NGSPEFKQKALAMDDYIKAGSL 179
QY 182 QTAPELERVLNFPQSNLCLKREKODESCSLTQALPSELKVSADCVSLTGAVALSMLETE 241

DB 180 KAGYEELVTNINIKDSQK-----KTDKCNLDSQKSNFTIENDKEPGVSGPLKINSADV 235
QY 242 IFLLQAOAGMP--EPGWRITDSHQWNTLLSLHNAQFDLQRTPEVARSATPLDLIKT 299
DB 246 AIDLQYGGFPADQVAMGLVDTPEKCKGLNTLKNAYQSTLFTPKIIAKVNAFILNLYIDK 295
QY 300 ALTHPPOKQAYGVTLPSVLFIAHDHNTNLNGALELN-WTLPGQPDNTPGEGELIAYE 358
DB 296 GFVSVDKGETA-----KFIPLVGHDSNLASLMSAMDPEKPYQLAQYEHPTIGGLVFO 348
QY 359 RWRRLSDNSQMIQVSLVFQTLQMRDTPPLSLNTPPEGEVLTLAGCE-ERNAQGCISLAG 417
DB 349 RWTQGTQKQDPMKVEYVYQTDQLRDNAYLSLETPPKVITLELKCPCPDKN--GICSMED 406
QY 418 FTQIVNEA 425
DB 407 FQKVMKA 414

RESULT 10

B87316
periplasmic phosphoanhydride phosphohydrolase [imported] - Caulobacter crescentus
C.Species: Caulobacter crescentus
C.Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C.Accession: B87316
R.Nierman, W.C.; Feldlyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laud, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolton
n, J.; Esmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A.Title: Complete Genome Sequence of Caulobacter crescentus.
A.Reference number: B87249; MUID:21173698; PMID:11259647
A.Accession: B87316
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-414 <STO>
A.Cross-references: UNIPROT:Q9A04; UNIPARC:UPI00000C70A0; GB:AB005673; NID:gl3421731; P
A.Genetics:
A.Gene: CC0542

Query Match 23.7%; Score 536; DB 2; Length 414;
Best Local Similarity 33.6%; Pred. No. 3, 1e-35;
Matches 145; Conservative 62; Mismatches 186; Indels 38; Gaps 11;

QY 9 LSL---LILPTQSAFAQSEPELKLSVSVYSRHGVAPTKATQLMQDVTPDAMPPTVPL 65
DB 3 LSLIAVVTTLALASAGASAE--TLEKVIILSRHGVSAKSSPERLEASARPPRFEPV 60
QY 66 LGELTPRGSELIAVIGHYWRQRLVADGLLPKCCGPOGQVATADDERTRKTGEAPAG 125
DB 61 AGHLTARGETLVARMGDYRRHRYAAQGLKPGDC--ASVYAMAVNTQRTIATAKAYRET 117
QY 126 LAPDCAITVHTQADTSSPDLFNPPLKTYGCOLDNANVTDAILERAGGSIAADFTGHYQAF 185
DB 118 LAPGCPYVNTVGE-GNIDPMFEVYKAGIVADALARAANAAGVGGDLTAMSGASHQEA 176
QY 186 RELERVLNFPQSNLCLKREKODESCSLTQALPSELKV-----SADCVSLTGAVALS 236
DB 177 EQLDALIMQCDKGC-----PPAGGRVFDAPKPFVDGSELGLGGEPAFA 223
QY 237 SMLTEIFLLQAOAGMPERGWG-RITDSHQWNTLLSLHNAQFDLQRTPEVARSATPLLD 295
DB 224 SCVTESLIMAMADRDPAGLQWKSLEBALTRSFPLQAEFDELRLRTPPYVARTLAGHLAD 283
QY 296 LIKTALTPHPPOKQAYGVTLFTSVLFIAGHDNTNLNGALELNMTLPG--QPDNTPPEGE 354
DB 284 RLATVLR---DGAALIG-PVDARVITLAGHGTLASLGGILRMEMTLPGVQPNQIQGGA 339
QY 355 LVFERWRRLSDNSQMIQVSLVFQTLQMRDTPPLSLNTPPEGEVLTLAGCEERNAQGCIS 414
DB 340 LVFERWRRLSDNSQMIQVSLVFQTLQMRDTPPLSLNTPPEGEVLTLAGCEERNAQGCIS 414
QY 415 LAGFTQIVNEA 425

Db 399 LEDPETVVRGA 409

RESULT 11

hypothetical protein F13D11.1 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C/Accession: T16058

R/Pulton, L.

submitted to the EMBL Data Library, November 1995

A/Description: The sequence of C. elegans cosmid F13D11.

A/Reference number: S69020

A/Accession: T16058

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-416 <FNU>

A/Cross-references: UNIPROT:Q19390; UNIPARC:UPI000007AD42; EMBL:U40939; MID:g1072175; PI

C/Genetics:

A/Gene: CESP:F13D11.1

A/Introns: 23/3; 59/3; 88/3; 138/3; 202/3; 244/1; 270/3; 312/3; 349/3; 377/1

C/Superfamily: mammalian acid phosphatase

Query Match 6.1%; Score 137; DB 2; Length 416;

Best Local Similarity 21.7%; Pred. No. 0.0035;

Matches 99; Conservative 61; Mismatches 166; Indels 130; Gaps 21;

1 MKALLIPPLSLIPPTQSAFAQSEPELKESVIVSHRGVRAFPKATQLMQDYTPD-- 57

1 MRVLTLF--PLFPV--AFG-----KLFKVOIWFHGRTRGHVLYFPFGDLDLVNDY 48

58 ---AMPPTVVKLGELTPRGELIAYIGHYWRQL-----VADGLLPKCGCPQSGQVAI 107

49 QQIAMP-----GELTRGIIIEEFGQLG---QRLKITYGHEFGDTYQPR-----DPHV 91

108 IADYDERTRTKGEAFAAGADPCATVHTQADTSSPDLFNPILKTGVCQLDNANVTDAIL 167

92 YTGKDNRTSASQAQMFAGFLPP-----NEDQTNVYELKMQPV-----AQUTDSI 136

168 ERAGGSITADFTGHYQTAARE-----LERYINFPQSNCLKREKQDSGSLTQALSEL 220

137 DWVSLGALDNCVYGEAQKRSSEYAEVMDQEKYDAELLQVRRHABE----- 184

221 KVSADCVSLTGAVSLASMLTEI---FLTQAQGMPEPGRITDSHQWNTLLSLHNAQFD 277

185 -----PIVAVKKNHYIDSLKRYITIQDDR-LPYPMAGYENRIILMSFLIHDA--- 233

278 LLQRTPEVARSRAATPLD---LTKTALTTPHPQQAQAVTLPSTVLEFIAGHDITNLANG 333

234 -----VVKVQNDSEVGDYHNELVMSYFETHQKNSKTKGV-----FISGHDITNLVTIW 279

334 GALLINWTLPGQDPNTTP--GGELVFERWRRLSDNSQMLQVSLVFTQLQGMKDTPLSLANT 392

280 ESLRID---GHPEDIPYGAHIAIE-----MHBPVQSLIKKFLSKGF 319

393 PGGEVKLTLAGCEERNAQMGSLAGFTQIVNEARIP 428

320 NOTVEILPHFC-SRSQNNDCTWDEFQRLVKKSRKP 354

RESULT 12

acid phosphatase (EC 3.1.3.2) ACP2 precursor [validated] - human

N/Alternate names: acid phosphatase; lysosomal

C/Species: Homo sapiens (man)

C/Date: 04-Dec-1992 #sequence_revision 01-Dec-1995 #text_change 09-Jul-2004

C/Accession: S06167; S05525; S01155

R/von Figura, K.

submitted to the EMBL Data Library, June 1989

A/Reference number: S06167

A/Accession: S06167

A/Molecule type: DNA

A/Residues: 1-423 <VON>

A/Cross-references: UNIPROT:P11117; UNIPARC:UPI0000131FE1; EMBL:X15525; MID:g34239; PIDN
R/Geter, C.; von Figura, K.; Pohlmann, R.
Eur. J. Biochem. 183; 611-616, 1989

A/Title: Structure of the human lysosomal acid phosphatase gene.

A/Reference number: S05525; MUID:89377828; PMID:2776754

A/Accession: S05525

A/Molecule type: DNA

A/Residues: 1-29 <GNI>

A/Cross-references: UNIPARC:UPI00001727CD

R/Pohlmann, R.; Krentler, C.; Schmidt, B.; Schroeder, W.; Lorkowski, G.; Culley, J.; Meri

EMBO J. 7; 2343-2350, 1988

A/Title: Human lysosomal acid phosphatase: cloning, expression and chromosomal assignment

A/Reference number: S01155; MUID:89052645; PMID:3191910

A/Accession: S01155

A/Molecule type: mRNA

A/Residues: 1-423 <POH>

A/Cross-references: UNIPARC:UPI0000131FE1; EMBL:X12548; MID:g34262; PIDN:CAA31064.1; PID

A/Note: part of this sequence, including the amino end of the mature protein, was confir

C/Genetics:

A/Gene: GDB:ACP2

A/Cross-references: GDB:118963; OMIM:171650

A/Map position: 11p11.2-11p11.11

A/Introns: 38/3; 70/3; 99/3; 150/3; 183/3; 213/3; 258/1; 285/3; 321/2; 380/1

C/Function:

A/Description: catalyzes the hydrolysis of a wide range of phosphate esters

C/Superfamily: mammalian acid phosphatase

C/Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydrolase

F/1-30/Domain: signal sequence #status predicted <SIG>

F/31-423/Product: acid phosphatase ACP2 #status experimental <MAT>

F/42/Active site: Arg #status predicted

F/92-133,167,177,191,267,322,331/Binding site: carboxylate (Asn) (covalent) #status pre

F/159-370,212-310,345-349/disulfide bonds: #status predicted

Query Match 5.8%; Score 132; DB 1; Length 423;

Best Local Similarity 23.3%; Pred. No. 0.0091;

Matches 97; Conservative 43; Mismatches 173; Indels 104; Gaps 20;

30 LESVIVSRHGVAAPFKATQLMQDYTPDAPWTPVVKLGELTPRGELIAYIGHYWRQLV 89

33 LRFVTLVLRHGDSVPVKT---YKDPYQEEBWPQGFQULKEGMLQHMLGQALKRKY- 87

90 ADGLLPKCGCPQSGQVAI1IADV-ERTRTKGEAFAAGADPCATVHTQADTSSPDL-- 146

88 -HGFL--NTSYHQEYVAVSTDFRTLMGAENLGLFP-----PNCMQR 129

147 FNP-LKTGVCQLDNANVTDAILERAGGSITADFTGHYQTAARELERYINFPQSNCLKREK 205

130 FNPISWQPIPVHTVPITE-----DRLKFPAGCPRYEQL 165

206 QDSGSLTQALPSELKVSADCVSLTGAVSLASMLTEIFL-----LQAQGMPEP 254

166 QNE---TRQTPRYQNESSRNAQFLDMVANETGLITDLETWNVYDTLFCQTHGLRLP 221

255 GWRGITDSHQWNTLLSLHNAQFDL---QRTPEVARSRAATPLDLIKTALTTPHPQQAQV 311

222 PW---ASPQMGQLSRKIDPSFRFLFGIQAERAKLQGGVLLAQIRKNTLMAATTSQ-- 276

312 GVTLPSTVLEFIAGHDITNLANGALBELANWTLPGQDPNTTPRGELVFERWRRLSDNSMIQ 371

277 ---LP-KLVYSADHTTLVALQWALDLY-----NGEAPAFYASCHIFELVYEDSGN---FS 324

372 VSLVFTQLQGMKDTPLSLANTPGGEVKLTLAGCEERNAQMGSLAGFTQIVNEARIP 428

325 VEMTF--RNESDKAPWPLSLP-----GCPHR-----CPLODFLR--TEPVVP 363

RESULT 13

acid phosphatase (EC 3.1.3.2) precursor - rat

C/Species: Rattus norvegicus (Norway rat)

C/Date: 29-Jan-1990 #sequence_revision 29-Jan-1990 #text_change 09-Jul-2004

C/Accession: A33395

R.Himeno, M.; Fujita, H.; Noguchi, Y.; Kono, A.; Kato, K.

Biochem. Biophys. Res. Commun. 162, 1044-1053, 1989

A:Title: Isolation and sequencing of a cDNA clone encoding acid phosphatase in rat liver

A:Reference number: A33395

A:Accession: A33395

A:Status: preliminary

A:Accession: A33395

A:Residues: 1-423 <HM>

A:Cross-references: UNIPROT:P20611, UNIPARC:UPI0000111FE2, GB:M27893, NID:g202933, PIND:

C:Superfamily: mammalian acid phosphatase

C:Keywords: phosphatidyl, phosphoprotein; phosphoric monoester hydrolase

F:1/Active site: Arg (phosphatidyl intermediate) #status predicted

F:42/Active site: His (phosphatidyl intermediate) #status predicted

Query Match 5.2%; Score 116.5; DB 2; Length 423;

Best Local Similarity 22.8%; Pred. No. 0.16;

Matches 102; Conservative 46; Mismatches 185; Indels 115; Gaps 22;

2 KALLIPFL--SLILPPLTQSAFAQSEPELKESVIVSRHGVAPRTKATQMLQDVTPEAM 59

10 QALLIQFLQMGCLMMPPIQAR-----RSLRFTLLYRHGDRSPVKA---YPKDPYGE 58

60 PTWPVYKGLTTPRGELIAYLGHYWRQRLVADGLLPKCCGQSGQVAIADVD-ERTKRT 118

59 EKMPOGFGQLTKEGMLQHWELGQALRQRY--HGFL--NASHRQEVYVRSSTDPEKTLMS 113

119 GGAFAAGLAPDCAITVHTQADTSSPDLFNLKTVGCQLDNNVTDAILERAGSIADEFT 178

114 AEAANLAGLFPF-----TEVQHNPNISWQPI-----PVHTVPITE----- 148

179 GHYQTAFLERLNVNPNQNLCKREKODESCSLQALPSELKVSADCVSLTGAVSLASM 238

149 -----DKLKFPLGFCRRYRQLQNE---TRQTEYQNMSTQNAQPLDMVANETG 194

239 LTRIFL-----LQAOQMEPEPGWGRITDSHQNNTLLSLNAQF-----DLIQ 280

195 LNNLTLETIMVYDTLPFCQTHGLLPPF--ASPTQVALSQLQDFSLPLFGIHDOYQ 251

281 RTPEVARSBATPLDLITKALTTPHPPOKQAVVTLPTSVLFIAGHDNTLANIGALELNM 340

252 K-----ARLQGVTLAQILKNTLMTATTSQF-----PKLLVYSADHTLVALLQMLNLY- 300

341 TLPGQPDNTPPGELVFERMRRLSDNSQWISLVFQTLQQRDTPLSINTPPEVLT 400

301 ----NGKQAPVASCIFELTYQ--EDNGNP-SVEMTFRDSK--KAPW-----PLT 341

401 LAGCEBNAQMGCSLAFPTQIVNEARIP 428

342 LFGCPHR-----CPLQDFLRLL-TEPVIP 363

Db

Qy

Db

Qy

Db

Qy

Db

A:Cross-references: FlyBase:FBgn0000032

C:Superfamily: mammalian acid phosphatase

C:Keywords: phosphoric monoester hydrolase

F:1-33/Domain: signal sequence #status predicted <SIG>

F:34-438/Product: acid phosphatase Acph-1 #status predicted <MNT>

Query Match 5.0%; Score 112.5; DB 2; Length 438;

Best Local Similarity 20.3%; Pred. No. 0.35;

Matches 82; Conservative 59; Mismatches 145; Indels 117; Gaps 19;

29 KLESVIVSRHGVAPRTKATQMLQDVTPEAMPT-----WVPLGELTPRGELIAYL 80

47 QAKFVHIVIRHGDRTFV-----DPTVTDWGRDKTFPTGWDITNLGQEHDL 95

81 GHYWRQRLVADGLLPKCCGQSGQVAIADVD-ERTKRTGEAFAGL-----APDC----- 130

96 GKMLNRY--SNLLPPI--YSNENIYQSTVDRTLSAQSNLAGLVEPQEDILMNDI 150

131 ---ATVHTQADTSSPDLFNLKTVGCQLDNNVTDAILERAGSIADEFTGHYQTAPE 187

151 NMQPIPIHTSPEREDP-----ILAAKAPCPAYDY-----E 180

188 LERVNPNQSNLCKREKODESCSLQALPSELKVSADCVSLTGAVSLASMTEIFLQ 247

181 LASLSSSEPRK-ALTEKRNLPAYISEKGRPVKTFID-----AQVINTLFTEN 229

248 AQGMEPEPGWGRITDSHQNNTLLSLNAQFDDLQRTPEVARSATPL-DLITKALTTPHP 306

230 LYNNLTLPKTKKVVYGRBELTVYS--NPAFASISYTRKLARLAKGLLMDIFGRF----- 281

307 QKQAVGTLTPSVLPI-AGHDNTLANIGALELNTLPGQPDNTPPGELVFERMRRLSD 365

282 KEKSSGLKPRSMVWVYSAHDITVAVSVAIAKTL-FEL-----HSPF----- 321

366 NSQWISLVFQTLQQRDTPPL-----SLNTPPEVLTLLAGC 404

322 ----YTAQIMELKRVDETNPPLVSIYKNTTAEPLDIPREC 359

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

A:Cross-references: FlyBase:FBgn0000032

C:Superfamily: mammalian acid phosphatase

C:Keywords: phosphoric monoester hydrolase

F:1-33/Domain: signal sequence #status predicted <SIG>

F:34-438/Product: acid phosphatase Acph-1 #status predicted <MNT>

Query Match 4.9%; Score 111.5; DB 2; Length 421;

Best Local Similarity 21.4%; Pred. No. 0.4;

Matches 98; Conservative 53; Mismatches 171; Indels 137; Gaps 23;

2 KALLIPF-----SLILPPLTQSAFAQSEPELKESVIVSRHGVAPRTKATQMLQDVTPE 56

8 QALLIQFLQMGCLMMPPIQARS-----LRFVTLVLRHGDERSVKT-----YKDP 53

57 DAWPTVPKGLTTPRGELIAYLGHYWRQRLVADGLLPKCCGQSGQVAIADVD-ERT 115

54 YQEKMPGFGQLTKEGMLQHWELGQALRQRY--HGFL--NASHRQEVYVRSSTDPEK 108

116 RKTGEAFAGLAPDCAITVHTQADTSSPDLFNLKTVGCQLDNNVTDAILERAGSIA 175

109 LMSAANLAGLFPF-----NEVQHFSPNISMQPI-----PVHTVPITE----- 146

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

```

Qy 176 DFTGHYQTAFRELERVLNFPQSNLCLKREKODESCSLTQALPSELKVADCVSLTGAVSL 235
Db 147 -----DRLKFPUGPCPRYEQLONE-----TROTPYQNSIQNAQFLNMYAN 189
Qy 236 ASMLTEIFL-----LQQAQGMPEPGWG-----RITDSHOWN--TLLSLHNAQF 276
Db 190 ETGLTNVTLERTIMNVYDTLFCEQTHGILLPWPASPOVQRLSOLKDFSLFLPGIHE--- 246
Qy 277 DLTORTPEVNASR-----ATPLDLIKTALTPHPQOKAYGVTLPTSVPFIAGHDTNL 329
Db 247 -----QVQKARLQGGVLLAQILKNLTLMATTSQFPK-----LLVYSAHDTTL 288
Qy 330 ANLGGALELNMWTLPGQPDNTPPGSELVPERWRRLSDNSQWIOVSLVFOTLQOMRDKTPPLS 389
Db 289 VALQMALNVY-----NGKQAPYASCHITELYQ--EDNGNF--SVEMTFRNDK---KAPWP 337
Qy 390 LNTPPGEVKLTLAGCCEERNNAQMCSLAGFTQIVNEARIP 428
Db 338 LILP-----GCPHR-----CPLQDFLRL-TBPVIP 361

```

Search completed: June 13, 2006, 10:29:30
 Job time : 26.2661 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: June 14, 2006, 10:58:44 ; Search time 5160.22 Seconds
(without alignments)
8030.264 Million cell updates/sec

Title: US-10-601-319-10
Perfect score: 2258
Sequence: 1 MKAILIPPLSLILPLTPQSA.....CSLAGTQIVNARIPACSL 432

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6366136 seqs, 31973710525 residues
Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abs/ABSSWB.spool/US10601319/runat_12062006_180050_27480/app.query.fasta_1
-DB=genembl -QFMT=fastcap -SUFFIX=rge -MINMATCH=0.1 -LOOFC=0 -LOOEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human4.0.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=ad8806h
-USER=US10601319 @CGN_1_1.3991 @runat_12062006_180050_27480 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_env:*
2: gb_pat:*
3: gb_ph:*
4: gb_pl:*
5: gb_pr:*
6: gb_ro:*
7: gb_scs:*
8: gb_sy:*
9: gb_un:*
10: gb_vl:*
11: gb_ov:*
12: gb_neg:*
13: gb_in:*
14: gb_om:*
15: gb_da:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2182	96.6	1901	2	AR636190 Sequence
2	2182	96.6	1901	2	AR636191 Sequence
3	2182	96.6	1901	2	AX356572 Sequence

4	2182	96.6	1901	2	AX356574
5	2182	96.6	1901	15	ECOARPPA
6	2182	96.6	3470	2	AX042376
7	2182	96.6	4060	2	AX042374
8	2182	96.6	6116	2	AX042375
9	2182	96.6	6708	2	AX042373
10	2182	96.6	17732	2	AX042378
11	2182	96.6	20623	2	AX042372
12	2182	96.6	110000	15	UD0086_10
13	2182	96.6	110000	15	AP009048_10
14	2182	96.6	110000	15	CP000038_10
15	2181	96.6	1901	2	AR636188
16	2181	96.6	1901	2	AX356570
17	2178	96.5	1315	15	AF537219
18	2176	96.4	1901	15	ECOACPHA
19	2176	96.4	1901	15	ECOACPHB
20	2176	96.4	1901	15	ECOACPHD
21	2176	96.4	1901	15	ECOACPHE
22	2175	96.3	1901	2	AR636189
23	2175	96.3	1901	2	AX356571
24	2175	96.3	1901	15	ECOACPHB
25	2172	96.2	1901	15	ECOACPHF
26	2171	96.1	1323	2	AR108133
27	2171	96.1	1323	2	AR127818
28	2171	96.1	1323	2	AR130956
29	2171	96.1	1323	2	AR492966
30	2171	96.1	1323	2	AR636185
31	2171	96.1	1323	2	AX052713
32	2171	96.1	1323	2	AX338966
33	2171	96.1	1323	2	AX356566
34	2164	95.4	1486	2	AR630587
35	2154	95.4	110000	15	CP000036_22
36	2152	95.3	5421	2	AX042377
37	2149	95.2	110000	15	AB005674_10
38	2147	95.1	1489	2	AR277753
39	2147	95.1	1489	15	AY496073
40	2145	95.0	110000	15	AB014073_10
41	2144	95.0	1486	2	AR630588
42	2144	95.0	110000	15	AB005174_12
43	2144	95.0	110000	15	AB005174_13
44	2144	95.0	110000	15	BA000007_12
45	2143	94.9	110000	15	AE014075_10

ALIGNMENTS

RESULT 1
LOCUS AR636190 1901 bp DNA linear PAT 20-APR-2005
DEFINITION Sequence 7 from patent US 6855365.
ACCESSION AR636190
VERSION AR636190.1 GI:62768873
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1901)
AUTHORS Short, J.M., Kretz, K.A., Gray, K.A., Barton, N.R., Garrett, J.B.,
O'Donoghue, E., and Mathur, R.J.
TITLE Recombinant bacterial phylases and uses thereof
JOURNAL Patent: US 6855365-A 7 15-FEB-2005;
Diversa Corporation; San Diego, CA
FEATURES
source
1..1901
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores: 3.74e-196 Length: 1901
Pred. No.: 2182.00 Matches: 424
Score: 98.1% Conservative: 0
Best Local Similarity: 98.1% Mismatches: 8

Query Match: 96.6% Indels: 0
 DB: 2 Gaps: 0
 US-10-601-319-10 (1-432) x AR636190 (1-1901)

```

Qy      1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db      188 ATGAAGAGGATCTTAATCCCATTTTATCTCTCGATTCCGTTAAACCCCGCAATCTGCA 247
Qy      21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValIleValSerArgHisGly 40
Db      248 TTCGCTCAGAGTGAAGCCGAGCTGAAGCTGGAAGATGGTGATGTCAGTCGTCATAGT 307
Qy      41 ValArgAlaProThrIleValAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
Db      308 GTGGGTGCTCCACCAAGGCGCACGCACTGATGCAAGATGTCACCCGACGACATGGCCA 367
Qy      61 ThrTrpProValLysLeuGlyGluLeuThrProArgGlyGlyIleValIleAlaTyrLeu 80
Db      368 ACCTGGCCGGGTAAACTGGGTGGCTGACACCGCGNGGTGGTGAAGCTAATGCCCTATCTC 427
Qy      81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
Db      428 GGACATTACCAACGCCAGCGCTGGGTAGCCGAGATTGCTGGCGAAAAAGGGCTGCCCG 487
Qy      101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgIleThrGlyGlu 120
Db      488 CAGTGTGATGAGGTGCGCAATTATTCCTGATGTCAGAGCGTACCCGTTAAACAGGCGAA 547
Qy      121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db      548 GCCTTCGCCGCCGGCTGGCACTGACCTGCAATACCGTACATACCAGGCAAGATACG 607
Qy      141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db      608 TCCAGTCCCGATCCCTTTTAAATCCTTAAAACTGGCGTTTGCCAACTGATTAACCG 667
Qy      161 AsnValThrAspAlaIleLeuGlnArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db      668 AACGTGACTGAAGCATCTCTACGACGGGAGGAGGAGGCTCAATGCTGACTTTAACGGG 727
Qy      181 TyrGlnThrAlaPheArgGluLeuGlnArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db      728 CGGCAAAAGCGGTTTCGGGAAGTGAACGGGTCTTAATTTCCGCATCAAACTGTGCG 787
Qy      201 LeuLysArgGluLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db      788 CTTAAACCTGAGAAACAGACGAAAGCTGTTCAATTACGACGAGCATTAACGGAACCTC 847
Qy      221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db      848 AAGGGACCGCGCAATGTCATTAACCGGTGGGTGAAGCTTCGCAATCAATGTGACG 907
Qy      241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
Db      908 GAGATATTTCTCTCTGCAACAGCAAGGGAATCCGCGAGCCGGGTGGGAAAGATCAC 967
Qy      261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
Db      968 GATTCAACACAGATGAAACACCTTGCTAAGTTGCAATACGCGCAATTTTATTTGCTACA 1027
Qy      281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleGlySerThrAla 300
Db      1028 CGCAGCGCAAGAGTGGCCGCGACCGCGCACCCGTTATTAAGTTTAATCAAGACACG 1087
Qy      301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db      1088 TTGAGCGCCCATCCACCGCAAAAACAGGCGTATGTGTGACATTACCACTTCAGTGTG 1147
Qy      321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTrp 340
Db      1148 TTTATCGCCGAGCAAGATACTAATCTGGCAATATCTGGCGGCGCACTGGAAGCTCAACTG 1207
  
```

```

Qy      341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db      1208 ACCTTCCCGGTACGCCGATTAACGCCCGCAGGTGTGTAACTGTGTGTAACGCTCG 1267
Qy      361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db      1268 CGTCGGCTAAGCGATTAACAGCGAGTTCAGGTTTCGCTGCTTCACAGACTTACAG 1327
Qy      381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
Db      1328 CAGATGGGTATTAACCGCGCTGTCAATTAATACGCGCGCGAGAGGTAAACTGACC 1387
Qy      401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db      1388 CTGGCAGATGTGAAGCGCAATGCCAGGCGCATGTGTCTTGTCAGAGTTTACGCA 1447
Qy      421 IleValAsnGlnAlaArgIleProAlaCysSerLeu 432
Db      1448 ATCGTAATGAAGCACCATACCGCGGTGCAATTG 1483

RESULT 2
AR636191
LOCUS AR636191 1901 bp DNA linear PAT 20-APR-2005
DEFINITION Sequence 9 from patent US 6855365.
ACCESSION AR636191
VERSION AR636191.1 GI:62768874
KEYWORDS
SOURCE
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 1901)
AUTHORS Short,J.M., Krezs,K.A., Gray,K.A., Barton,N.R., Garrett,J.B.,
O'Donoghue,E. and Mathur,E.J.
TITLE Recombinant bacterial phycases and uses thereof
JOURNAL Patent: US 6855365-A 9 15-FEB-2005,
DIVERSA Corporation; San Diego, CA
FEATURES
source
1..1901
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 3,74e-196 Length: 1901
Score: 2182.00 Matches: 424
Percent Similarity: 98.1% Conserves: 0
Best Local Similarity: 98.1% Mismatches: 8
Query Match: 96.6% Indels: 0
DB: Gaps: 0

US-10-601-319-10 (1-432) x AR636191 (1-1901)
Qy      1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db      188 ATGAAGAGGATCTTAATCCCATTTTATCTCTCGATTCCGTTAAACCCCGCAATCTGCA 247
Qy      21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValIleValSerArgHisGly 40
Db      248 TTCGCTCAGAGTGAAGCCGAGCTGAAGCTGGAAGATGGTGATGTCAGTCGTCATAGT 307
Qy      41 ValArgAlaProThrIleValAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
Db      308 GTGGGTGCTCCACCAAGGCGCACGCACTGATGCAAGATGTCACCCGACGACATGGCCA 367
Qy      61 ThrTrpProValLysLeuGlyGluLeuThrProArgGlyGlyIleValIleAlaTyrLeu 80
Db      368 ACCTGGCCGGGTAAACTGGGTGGCTGACACCGCGNGGTGGTGAAGCTAATGCCCTATCTC 427
Qy      81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
Db      428 GGACATTACCAACGCCAGCGCTGGGTAGCCGAGATTGCTGGCGAAAAAGGGCTGCCCG 487
Qy      101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgIleThrGlyGlu 120
  
```

```

Db      ||| 488 CAGTCTGGTCAGGTCCTCGATTATTCGTGATGTCAGACGAGGTACCCGTAAACAGCGCAA 547
Qy      ||| 121 AAlpheaAlaAGlyLeuAlaProApCyAlaAlIethrValHsthrglnAlaApThr 140
Db      ||| 548 GCGCTTCGCGCGCGGCTGACCTGACTGTCGCAATACCGTACATCCCGAGCAATACG 607
Qy      ||| 141 SerSerProApPProleuPheApnPProleuLeuThrglyValCyeglneuApAanla 160
Db      ||| 608 TCCAGTCCCGATCCGTATTATTCCTTAATAACCTGGCTTTGCCAATCGAATACGCG 667
Qy      ||| 161 AasnValThraPAlaAlIeuglnuArgAlaglyGlySerIleAlaApPheThrglyHs 180
Db      ||| 668 AACGTGACTGACGCGATCTCTCAGACGAGGAGGAGGTCAATTGCTGACTTACCGGCA 727
Qy      ||| 181 TyrGlnThraAlaPheArgrgluLeuglnuArgValIleuApnPheProglInSerAenleuCy 200
Db      ||| 728 CCGCAAAACCGCGCTTTCGCAACTGGAACGAGGTGCTTAATTTTCCGCAATCAACTGTGC 787
Qy      ||| 201 LeuLyArgrgluLyeglnuApGluSerCySerIleuThrglnAlaLeuProSerGluLeu 220
Db      ||| 788 CTTAAACGTGAGAAACAGAACGAACTGTTCACTTAACGAGCATTAACATCGAATCTC 847
Qy      ||| 221 LysValSerAlaApCyAlaSerIleuThrglyAlaValSerIleuAlaSerMetLeuThr 240
Db      ||| 848 AAGGTGAGCGCGCAATGCTCTCATTAACCGGTGCGTAAAGCTCGCATCAATGCTGACG 907
Qy      ||| 241 GlnIlePheLeuLeuGlnInAlaGlnGlyMetProGlnuProgluYrThrglyAlaThr 260
Db      ||| 908 GAGAAATATTTCTCTGTGCAACAGACAGGAAATGCCGAGCCGCGGTGGGAAAGATCACCC 967
Qy      ||| 261 AApSerHIsGlnITrPAnThrIleuLeuSerIleuHIsAAsnAlaGlnPheApPheLeuGln 280
Db      ||| 968 GATTACACACATGGAACACCTTGCTAATTTGCTATACCGCAATTTGTTGCTTACAA 1027
Qy      ||| 281 ArgThrProGlnuValAlaArgSerArgAlaThrProleuLeuApPheIleuYsThraAla 300
Db      ||| 1028 CGCAGCGCAGAGGTGCGCGAGCGCGCCACCCGTTATTGATTGATCAAGACAGCG 1087
Qy      ||| 301 LeuThrProHIsProProglInuYeglnAlaYrGlyValThleuPProThIsSerValLeu 320
Db      ||| 1088 TTGAGCGCCCATCCACCGCAAAACAGCGGTATGGTGACATTAACCCACTTCAGTGC 1147
Qy      ||| 321 PheIleAlaGlyHIsAAsPThraPheLeuAlaAenleuGlyAlaIleuGlnleuAenThr 340
Db      ||| 1148 TTTATCGCGGACAGATATCTAATCTGGCAATCTCGCGGCGCACTGAGGCTCAACTGC 1207
Qy      ||| 341 ThrIleuProgluYglnProApPAnThrProProgluYglnuLeuValPheGlnuArgTrp 360
Db      ||| 1208 ACGCTTCCCGGTCAAGCCGATTAACAGCGCGCAAGGTGATGTTGTAACGCTGC 1267
Qy      ||| 361 ArgAlaGluSerArApPAnSerGlnITrIleGlnValSerIleuValPheGlnThrLeuGln 380
Db      ||| 1268 CGTCCGCTTAAGGAAACAGCCAGTGATTCAGGTTTCCGTCTTCCAGACTTACAG 1327
Qy      ||| 381 GlnMetArApPArPLeuThraPProleuSerIleuAenThrProProgluGlnValuYsLeuThr 400
Db      ||| 1328 CAGATGCGGTATTAACCGCGCTGTCACTTAATACGCGCGCGAGAGGTGAACCTGACC 1387
Qy      ||| 401 LeuAlaGlyCyeglnuGlnuArgAanAlaGlnGlyMetCySerSerIleuAlaGlyPheThrgln 420
Db      ||| 1388 CTGGCAGGATGTGAAGAGGAATGCGACGCGCATGTGTTGCTTGGCAGGTTTACGCA 1447
Qy      ||| 421 IleValaenGlnuAlaArgIleProAlaCySerIleu 432
Db      ||| 1448 ATCGTAATGAAGCAGCATACCGCGGTGCAAGTTTG 1483

```

```

RESULT 3
LOCUS   AX356572      1901 bp      DNA      linear      PAT 06-FEB-2002
DEFINITION Sequence 7 from Patent WO0190333.
ACCESSION AX356572
VERSION   AX356572.1 GI:18621059

```

```

KEYWORDS
SOURCE   Escherichia coli
ORGANISM Escherichia coli
REFERENCE
AUTHORS  Short, J.M., Kretz, K.A., Gray, K.A., Barton, N.R., Garrett, J.B. and
          O'Donoghue, E.
TITLE    Recombinant bacterial phytases and uses thereof
JOURNAL  Patent: WO 0190333-A 7 29-NOV-2001;
          DIVERSA CORPORATION (US)
FEATURES
          source
            1..1901
              /organism="Escherichia coli"
              /mol_type="unassigned DNA"
              /db_xref="taxon:562"
ORIGIN
Alignment Scores:
Pred. No.:      3 74e-196
Score:          2182.00
Percent Similarity: 98.1%
Best Local Similarity: 98.1%
Query Match:    96.6%
DB:             2
              Gaps: 0
US-10-601-319-10 (1-432) x AX356572 (1-1901)
Qy      1 MetLyAlaAlaIleuIleProPheLeuSerIleuIleProleuThrProglInserAla 20
Db      188 ATGAAGAGCATCTTAATCCATTATTTATCTCTGATTCGTTAACCCGCAATGTGA 247
Qy      21 PheAlaGlnSerGlnuProgluLeuLySerValIleValSerArgHIsGly 40
Db      248 TTGCGCTCAGAGTGAACCGGAGCTGAAGCTGGAAGTGTGCTGATTTGCTGCTCATGCT 307
Qy      41 ValArgAlaProThrLyAlaAlaThrglnIleuMetGlnuApValThraPAsnAlaTrpPro 60
Db      308 GTGCGTGTCTCAACAGGACAGGCAAGTGAAGAGATGACCCCGAGACGATGCGCA 367
Qy      61 ThrTrpProValuYsLeuGlnuYglnuLeuThrProArgIlyGlyIleuIleAlaYrIleu 80
Db      368 ACTGCGCGGTAAACTGGGTGGCTGACACCGCGNGGTGTGACTAATGCCCTATCTC 427
Qy      81 GlyHIsYrTrpArgGlnuArgLeuValAlaApArgIleuLeuProLyCyeglyCyPro 100
Db      428 GGAACATTAACCAACGCGCAGCTGCTGTAACCGGAGATTTGCTGCGCAAAAGGCGTGC 487
Qy      101 GlnSerGlyGlnuValAlaAlaIleAlaPheValaApGlnuArgTrpArguYsThrglyGln 120
Db      488 CAGTCTGGTCAGGTGCGAATTAATGCTGATGTCGACGAGCGTACCCGTTAAACAGGCGCA 547
Qy      121 AAlpheaAlaAGlyLeuAlaProApCyAlaAlIethrValHsthrglnAlaApThr 140
Db      548 GCGTTCGCGCGCGGCTGCACTGATCTGGAATACCGTACATACCGCAAGCATACG 607
Qy      141 SerSerProApPProleuPheApnPProleuLeuThrglyValCyeglneuApAanla 160
Db      608 TCCAGTCCCGATCCGTATTATTCCTTAATAACCTGGCTTTGCCAATCGAATACGCG 667
Qy      161 AasnValThraPAlaAlIeuglnuArgAlaglyGlySerIleAlaApPheThrglyHs 180
Db      668 AACGTGACTGACGCGATCTCTCAGACGAGGAGGTAATTTGCTGACTTACCGGCA 727
Qy      181 TyrGlnThraAlaPheArgrgluLeuglnuArgValIleuApnPheProglInSerAenleuCy 200
Db      728 CCGCAAAACCGCGCTTTCGCAACTGGAACGAGGTGCTTAATTTTCCGCAATCAACTGTGC 787
Qy      201 LeuLyArgrgluLyeglnuApGluSerCySerIleuThrglnAlaLeuProSerGluLeu 220
Db      788 CTTAAACGTGAGAAACAGAACGAAAGCTGTTCACTTAACGAGCATTAACATCGGAATCTC 847
Qy      221 LysValSerAlaApCyAlaSerIleuThrglyAlaValSerIleuAlaSerMetLeuThr 240

```

Db	848	AAGGTGAGCGCGCAATGCTCTCAATTAACCGGTGGTGAACCTTCGCATTCATGCTGACG	907
Qy	241	GIUILEPHELEULENGINGINALAGINGLYMECPROGIUPROGIYTRPGIYARGI1ETHR	260
Db	908	GAGNATATTTCTCTCGACAAACGACACAGGAAATGCCGAGCCGGGTGGGAGAAGATCACG	967
Qy	261	ASPSETHISGINTTPAETHNTHLEULEUSERLEUHIASAALAGINPHEASPLEULEGIN	280
Db	968	GATTCACACCAAGTGAGAACACCTTGCTTAAGTTTGACATAACGGCAATTTTATTTCTCTCAA	1027
Qy	281	ARGTHPRPROGIUVALAIAARGSERARGIAATHRPROULEUAPLEUIILEYETHRALA	300
Db	1028	CGCAGCGCCAGAGGTGGCCCGCAGCCGCCACCCCGTTATTAAGATTGATCAAGACAGCG	1087
Qy	301	LEUTHRPROHISPROPROGINLUSGINALATYRGIVALTTHLEUPROTHSERVALLEU	320
Db	1088	TTGACGCCCATTCACCGCAAAAACAGCGCGTRHGTGTGACATTAACCATTCAGTGTGCTG	1147
Qy	321	PHIELLEALAGIYHIASAPTHRASNULIAASULENGIYGIYALALEUGIULEUASNTTP	340
Db	1148	TTTATTCGCGGACACAGTACTAATCTGCAATCTCGCGCGCGCACTGACTCACTGCG	1207
Qy	341	THRIEUPROGIYGINPRHAPASANTHRPROPROGIYGIYGIULEUVALPHEGIUAARGTTP	360
Db	1208	ACGCTTCCCGGTCAAGCCGGAATTAACCGCGCAAGTGTGTGAACGTGTTTGAAACGCTGG	1267
Qy	361	ARGATGLESERASPSASNSERGINTRPILEGINALSERLEUVALPHEGINTHREUGIN	380
Db	1268	CGTCGGCTCAACGATTAACGCCACGAGATTCAAGTTTGCTGTGCTTCCACATTTACAG	1327
Qy	381	GINMETAGSAPLYETHRPROLEUSERLEUANTHRPROPROGIYGIYUVALLYSLEUTHR	400
Db	1328	CAGATGCGTGATTAACACCGCCGCTGCATTTAATAATACGCCGCCGAGAGAGTGAAACTGACC	1387
Qy	401	LEUIAGIYCYGSIUGIUNARGASNAIAGINLYMETCYSSERLEUVALGLYPHETHGIN	420
Db	1388	CTGGCAGGATGTGAAGACGAAATGCGAGGCGCATGTGTTGTTGGCAGGTTTATTCAGCAA	1447
Qy	421	ILEVALAENGILUALAARGILEPROHIACYSSERLEU	432
Db	1448	ATCGTGATGAAGCACGCAATCCGGCGTGCACTTTG	1483
RESULT 4			
LOCUS	AX356574	1901 bp	DNA
DEFINITION	Sequence 9 from Patent WO0190333.		Linear
ACCESSION	AX356574		PAT 06-FEB-2002
VERSION	AX356574.1	GI:18621060	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS	1 Short, J.M., Kretz, K.A., Gray, K.A., Barton, N.R., Garrett, J.B. and O'Donoghue, E.		
TITLE	Recombinant bacterial phytases and uses thereof		
JOURNAL	Parent: WO 0190333-A 9 29-NOV-2001;		
FEATURES	DIVERSA CORPORATION (US)		
source	Location/Qualifiers		
	1..1901		
	/organism="Escherichia coli"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:562"		
ORIGIN			
Alignment Scores:			
Pred. No.:	3,74e-196	Length:	1901
Score:	2182.00	Matches:	424
Percent Similarity:	98.1%	Conservative:	0
Best Local Similarity:	98.1%	Mismatches:	8
Query Match:	96.6%	Indels:	0

[illegible]

Db 1208 AGCTTCCGGTACGCGGATACACGCCGCCAGGTGTGAACCTGTTGAACCTG 1267
Qy 361 ATGATGLeuSerAspAsnSerGIntPILegIlnValSerLeuValPheGlnThrLeuGln 380
Db 1268 CGTGGCTAAGCATTAACGACGAGTGAATTCAGATTTCGTGCTTCAGACTTTACAG 1327
Qy 381 GlnMetArgAspIleThrProLeuSerLeuAsnThrProProGlyGluValIleLeuThr 400
Db 1328 CAGATCGGTATTAACGCGCTGTCAATTAACGCGCCGAGAGGTGAACCTGACC 1387
Qy 401 LeuValAGLYCyseGluGluArgAsnAGlnGlyMetCyseSerLeuAGlyPheThrGln 420
Db 1388 CTGGCAGGATGTGAAGAAGCAATGCGCAGCGCATGTTCCTTCGACAGTTTACGCCAA 1447
Qy 421 ILeValAsnGluAlaArgIleProAlaCyseSerLeu 432
Db 1448 ATCGTGAATGAAGCAGCATACCGCGGTGCAGTTTG 1483
RESULT 5
ECOPAPPA 1901 bp DNA linear BCT 04-APR-2002
LOCUS Escherichia coli periplasmic phosphoanhydride phosphohydrolase
DEFINITION (appa) gene, complete cds.
ACCESSION M58708
VERSION M58708.1 GI:145283
KEYWORDS
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE 1 (bases 1 to 1901)
Dassa,J., March,C. and Boquet,P.L.
The complete nucleotide sequence of the Escherichia coli gene appA
reveals significant homology between pH 2.5 acid phosphatase and
glucose-1-phosphatase
J. Bacteriol. 172 (9), 5497-5500 (1990)
JOURNAL PUBLISHED 2 (sites)
REFERENCE 2168385
AUTHORS Ostani,K., Harms,E.H., Stevls,P.B., Kuciel,R., Zhou,M.M. and Van
Etten,R.L.
Overexpression, site-directed mutagenesis, and mechanism of
Escherichia coli acid phosphatase
J. Biol. Chem. 267 (32), 22830-22836 (1992)
JOURNAL PUBLISHED 3 (sites)
REFERENCE 1429631
AUTHORS Ostani,K. and Van Etten,R.L.
Asp304 of Escherichia coli acid phosphatase is involved in leaving
group protonation
U. Biol. Chem. 268 (28), 20778-20784 (1993)
JOURNAL PUBLISHED 8407904
FEATURES
source location/Qualifiers
1..1901 /organism="Escherichia coli"
/mol_type="genomic DNA"
/db_xref="taxon:562"
3..1522 /gene="appa"
3..8 /gene="appa"
16..108 /gene="appa"
/note="ORF"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AA072085.1"
/db_xref="GI:145283"
/translation="MWTLMFVGIILMCSLTVLVWLDPLIKS"
118..149
133..138 /gene="appa"
/gene="appa"
stem_loop
-35_signal
-35_signal
/gene="appa"

-10_signal 156..161
/gene="appa"
RBS 172..177
/gene="appa"
CDS 188..1486
/gene="appa"
/standard_name="acid phosphatase"
/codon_start=1
/transl_table=11
/product="periplasmic phosphoanhydride phosphohydrolase"
/protein_id="AA072086.1"
/db_xref="GI:145285"
/translation="MKAILIPLSILIPTPQAPAPQSPRLKTSVTVSRGVRAP
TKATQMDVTPDAMPFWKGMILTPRGELIATIGTQRQRLVADGILAKGCCQS
GQVATADVDETRRTGFAAGLAPDCATVHTQADTSSPDLFENPLKTVGCQDNA
NVTDAILSRAGSIADFTGHRQTARERLVNTPQSNCLKREKDESCSLTALPS
ELKVSADNVSLTGAVSLASMLTEIFLQQAQMPGMRITDSHOMTLLSLHNAQF
YLLQRTPEVARSRAVPLDLITKLTALPHPRQKQAVGVLPSTVLPFAGHDITLANIG
ALELWTLPGQPDNTPRGELVFERMRISDSQWTVSLVPTQLQMRDKTPLSINT
PRGEVKTLAGCEBNAQMCSLAGFTQIVNARIPLACSL"
188..253
/gene="appa"
254..1483
/gene="appa"
/product="periplasmic phosphoanhydride phosphohydrolase"
299..301
/gene="appa"
/standard_name="H303A mutant"
/note="created by site directed mutagenesis"
/citation=[2]
299..301
/gene="appa"
/standard_name="R16A mutant"
/note="created by site directed mutagenesis"
/citation=[2]
299..301
/gene="appa"
/standard_name="R20A mutant"
/note="created by site directed mutagenesis"
/citation=[2]
299..301
/gene="appa"
/standard_name="H17N mutant"
/note="created by site directed mutagenesis"
/citation=[2]
302..304
/gene="appa"
/standard_name="D304A mutant"
/note="created by site directed mutagenesis"
/citation=[3]
/phenotype="Asp 304 changed to Ala"
/replace="gct"
1163..1165
/gene="appa"
/standard_name="D304B mutant"
/note="created by site directed mutagenesis"
/citation=[3]
/phenotype="Asp 304 changed to Glu"
/replace="gag"
1497..1522
/gene="appa"
/note="tho-independent terminator"

ORIGIN
Alignment Scores:
Pred. No.: 3,74e-196 Length: 1901
Score: 2182.00 Matches: 424
Percent Similarity: 98.1% Conservative: 0

QY 101 GlnSerGlyValIleValIleIleValIleAspValAspGluArgIleThrGlyGlu 120
DB 2111 GAGTCTGGTCAAGGTCGGGATTTATGCTGATGTGACAGGGTACCGGTAAACAGGGGAA 2170
QY 121 AlaPheAlaIleValIleValIleProAspCysAlaIleThrValHisThrGlnIleAspThr 140
DB 2171 GCTTGGCCGGCGGGGTGGCACTGTGCAATTAACGTATACCCAGGCAAGTACG 2230
QY 141 SerSerProAspProLeuPheAsnProLeuValThrGlyValCysGlnIleuAspAsnAla 160
DB 2231 TCCAGTCCCGATCCGTTATTTATCTCTAAACATGGCGCTTGGCAACGTGATTAACGCG 2290
QY 161 AsnValIleAspAlaIleIleuGluArgAlaGlySerIleAlaAspPheThrGlyHis 180
DB 2291 AACGTACTGACCGCATCTCTACAGCGGCAAGAGGGTCATTTCTGACTTTACCGGCGCAT 2350
QY 181 TyrGlnThrAlaPheArgIleuGluArgValIleuAsnPheProGlnSerAsnLeuCys 200
DB 2351 CGGCAACGGCGCTTGGCGAATCGAAGCGGTCCTTAATTTTCCGCAATCAACTTGTGC 2410
QY 201 LeuValArgGluValGlnAspGlnSerCysSerLeuThrGlnIleuProSerGluLeu 220
DB 2411 CTTAAACGTGAGAAACAGGACGAAGCTGTTCAATTAACGCAATTAACATCGAATCTC 2470
QY 221 LysValSerIleAspCysValSerLeuThrGlyValIleValSerIleuAspSerLeuThr 240
DB 2471 AAGGTAGCGCCCAATGCTCTATTAAACGGTGGGTAAAGCTTCCCATCAAGTCGAGCG 2530
QY 241 GluIlePheLeuIleuGlnIleValIleGlnIleMetProGlnIleProGlnIleThr 260
DB 2531 GAGATTTTCTCTCTGCAACAGCAAGGATTCGCGAGCGGGGTGGGAGAGATCAAC 2590
QY 261 AspSerHisGlnIleThrAsnThrIleuLeuSerLeuHisAsnAlaGlnPheAspLeuGln 280
DB 2591 GATTCAACACAGGAGAAACATCTGTAAATTTGTCATTAACGCGCAATTTATTTGCTACAA 2650
QY 281 ArgThrProGlnValAlaIleArgSerAlaIleThrProLeuIleuAspIleIleThrAla 300
DB 2651 CGCAGCGCAAGAGTTCGCCGAGCGCGCACCCCGTTATGATTTGATCAAGACAGCG 2710
QY 301 LeuThrProHisAspProGlnIleValIleArgIleValIleThrLeuProThrSerValIle 320
DB 2711 TTGAGCGCCCATCCACCGCAAAACAGGGGTATGGTGTACATTAACCTTCAAGTCTG 2770
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeuGlnIleuAsnThr 340
DB 2771 TTTATCGCCGACACGATACTAATCTGCGAAATCTCGCGCGGCACTGAGCTCAACTG 2830
QY 341 ThrLeuProGlnValIleProAspAsnThrProProGlnIleValIleuValIleGluArgTrp 360
DB 2831 AGCGTTCGGGTCAAGCGGATTAACAGCGCGCAAGGTGGTGAACCTGTGTTGAAACGCTG 2890
QY 361 ArgArgLeuSerAspAsnSerGlnIleGlnIleValSerLeuValIlePheGlnThrLeuGln 380
DB 2891 CCTCGGCTAAGGATTAACGCAAGTGAATTCAGGTTTCGCTGCTTCCAGACTTTACG 2950
QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlnIleValIleuAsnThr 400
DB 2951 CAGATGCGGTAAACCGCGCTGTCAATTAATACCGCCCGGAGAGGTGAACACTGAC 3010
QY 401 LeuAlaGlyCysGlnIleuArgAsnAlaGlnIleMetCysSerLeuAlaGlyPheThrGln 420
DB 3011 CTGGCAGGATGTGAAGAGGAAATGGCAGGGGCAATGTGTTGTTGCAAGTTTACGCAA 3070
QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
DB 3071 ATCTGTAATGAAGCAGCATACCGCTTGCAGTTTG 3106
RESULT 7
AX042374 4060 bp DNA linear PART 23-NOV-2000
LOCUS AX042374
DEFINITION Sequence 3 from Patent WO0064247.

ACCESSION AX042374
VERSION AX042374.1 GI:11340992
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE
1 Foreberg, C.W., Golovan, S. and Phillips, J.P.
Transgenic animals expressing salivary proteins
Patent: WO 0064247-A 3 02-NOV-2000;
UNIVERSITY OF GUELPH (CA)
FEATURES
source location/Qualifiers
1. 4060
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="RIS/APPA + intron transgene"
ORIGIN
Alignment Scores:
Pred. No.: 1.14e-195 Length: 4060
Score: 2182.00 Matches: 424
Percent Similarity: 98.1% Conservative: 0
Best Local Similarity: 98.1% Mismatches: 8
Query Match: 96.6% Indels: 0
DB: 2 Gaps: 0
US-10-601-319-10 (1-432) x AX042374 (1-4060)
QY 1 MetLysAlaIleLeuIlePhePheLeuSerLeuIleProLeuThrProGlnSerAla 20
DB 1811 ATGAAGCCATCTTAATCCCATTTTATCTCTGATTCGTTAACCCGCAATCTGA 1870
QY 21 PheAlaGlnSerGluProGluIleuLysGlnSerValIleIleValSerArgHisGly 40
DB 1871 TTCGTCAGAGAGAGCGCGAGCTGAAGCTGAAGTGTGATGTCAGTGTACTGCT 1930
QY 41 ValAlaGlnAspThrIleValAlaThrGlnLeuMetGlnAspValIleThrProAspAlaTrpPro 60
DB 1931 GTGCGGTCTCCAAACAGGCAAGCCAGCAACGTAGCAGAGATGCAACCCAGACCATGGCCA 1990
QY 61 ThrTrpProValIleLeuGlnIleuLeuThrProAspGlyGlyGlnLeuIleAlaTyrLeu 80
DB 1991 ACCTGGCCGCTAAACCTGGGTTGCTGTACACCGCGGGTGGTGAATTCGCTATCTTC 2050
QY 81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
DB 2051 GGAATTAACCAAGCGCAAGGCTGTGGTACGCAAGGATTCGCGAAAGGCGTCCCG 2110
QY 101 GlnSerGlyValIleValIleIleAlaAspValAspGluArgTrpThrArgIleThrGlyGlu 120
DB 2111 GAGTCTGGTCAAGGTCGGGATTTATGCTGATGTGACAGGGTACCGGTAAACAGGGGAA 2170
QY 121 AlaPheAlaIleValIleValIleProAspCysAlaIleThrValHisThrGlnIleAspThr 140
DB 2171 GCTTGGCCGGCGGGGTGGCACTGTGCAATTAACGTATACCCAGGCAAGTACG 2230
QY 141 SerSerProAspProLeuPheAsnProLeuValThrGlyValCysGlnIleuAspAsnAla 160
DB 2231 TCCAGTCCCGATCCGTTATTTATCTCTAAACATGGCGCTTGGCAACGTGATTAACGCG 2290
QY 161 AsnValIleAspAlaIleIleuGluArgAlaGlySerIleAlaAspPheThrGlyHis 180
DB 2291 AACGTACTGACCGCATCTCTACAGCGGCAAGAGGGTCATTTCTGACTTTACCGGCGCAT 2350
QY 181 TyrGlnThrAlaPheArgIleuGluArgValIleuAsnPheProGlnSerAsnLeuCys 200
DB 2351 CGGCAACGGCGCTTGGCGAATCGAAGCGGTCCTTAATTTTCCGCAATCAACTTGTGC 2410
QY 201 LeuValArgGluValGlnAspGlnSerCysSerLeuThrGlnIleuProSerGluLeu 220
DB 2411 CTTAAACGTGAGAAACAGGACGAAGCTGTTCAATTAACGCAATTAACATCGAATCTC 2470

QY	221	lyseValser1aaepCyvsValserLeuthrGlyAlaValserLeuAlaSerMetLeuthr	240
Db	2471	AAGGTAGACGGCGCAAAATGTCTCATTAACCGGTGGGTAAACCTTCGCATCAATGCTGACG	2530
QY	241	GluliePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr	260
Db	2551	CAGATATTCTTCCTCCACCAACGACACAGGGAATGCCGAGGCCGGGTGGGAGAGATCACCC	2590
QY	261	AspSerHisGlnIleThrAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln	280
Db	2591	GATTCAACAACAGTGAACAACCTTGCTAAAGTTTGACATAACGGCAATTTTATTTGCTACAA	2650
QY	261	ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLeuThrAla	300
Db	2651	CGCAAGCCACAGAGTTGCCCGCAGCCGCCACCCCGTTATTAGATTGATCAAGACAGCG	2710
Db	2711	LeuThrProHisAspProGluValGlnAlaTyrGlyValThrLeuProThrSerValLeu	320
QY	321	PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaAlaGluIleuAsnThr	340
Db	2771	TTTATCGCCGACACAGATATCTATGCGAAATCTCGCGCGGCGACCTGGAAGCTACACG	2830
QY	341	ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp	360
Db	2831	ACGCTTCCCGGTCAACCGGATTAACGCCGCCAGCGGTGTGAATCTGTGTTGAACGCTGG	2890
QY	361	ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln	380
Db	2891	CGTCGGCTAAAGGATTAACGACGAGGATACAGGTTGCTGTGCTTCACAGCTTTACAG	2950
QY	381	GlnMetArgAspGlyThrProLeuSerLeuAsnThrProProGlyGlyValLeuLeuThr	400
Db	2951	CAGATCGCTGATTAACACCGCGCTGCATTAAATAACGCCGCCCGACGAGAGTAAACTGACC	3010
QY	401	LeuAlaGlyCyseGluGluArgAsnAlaGlnGlyMetCyseSerLeuAlaGlyPheThrGln	420
Db	3011	CTGGCAGGATGTGAAGACGCAAAATCGCAGGGCATGTGCTGGTCAGGTTTACGCCAA	3070
QY	421	IleValAsnGluAlaArgIleProAlaCyseSerLeu	432
Db	3071	ATCGTAATGAAGACAGCATACCGCTTGACAGTTTG	3106
RESULT 8			
LOCUS	AX042375	6116 bp	DNA linear PAT 23-NOV-2000
DEFINITION	Sequence 4 from Patent WO0064247.		
ACCESSION	AX042375		
VERSION	AX042375.1	GI:11340993	
KEYWORDS	.		
SOURCE	synthetic construct		
ORGANISM	other sequences; artificial sequences.		
AUTHORS	1		
TITLE	Foreberg,C.W., Golovan,S. and Phillips,J.P.		
JOURNAL	Transgenic animals expressing salivary proteins		
FEATURES	Patent: WO 0064247-A 4 02-NOV-2000;		
source	UNIVERSITY OF GUELPH (CA)		
	Location/Qualifiers		
	1..6116		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:33630"		
	/note="R15/AppA plasmid with pLICAT3 vector"		
ORIGIN			
Alignment Scores:			
Pred. No.:	2,07e-195	Length:	6116
Score:	2182.00	Matches:	424
Percent Similarity:	98.1%	Conservative:	0
Best Local Similarity:	98.1%	Mismatches:	8
Query Match:	96.6%	Indels:	0

DB:	2	Gaps:	0
US-10-601-319-10 (1-432) x AX042375 (1-6116)			
QY	1	Methylsalalileu1leProPheuserSerleu1leProleuThrProGlnSerAla	20
Db	1811	ATGAAGGCATCTTAATCCATTTTATCTCTTCGATTCCGTTAACCCCGCATCTGCA	1870
QY	21	PhealaginSerGluProGluLeuLeuLysLeuSerValValIleValSerArgHisGly	40
Db	1871	TTCCGCTCAGAGTGAACCGGAGCTGAAGCTGAAAGTGGTGAATGTTCAGTCGCATAGT	1930
QY	41	ValAlaGlaProThrThyValAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro	60
Db	1931	GTGCGTGTCTCCAAACCAAGGCCACGCACATGTATGCAAGATGTCACCCAGACGATGGCA	1990
QY	61	ThrTrpProValLysLeuGluGluLeuThrProArgGlyGlyGluLeuIleAlaTyLeu	80
Db	1991	ACCTGACCGGTAAACCTGGGTGGCTTGACACCGCGCGGTGTGACTAATGGCTATCTC	2056
QY	81	GlyHisTyTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro	100
Db	2051	GGACATTCACAAACGCCCAAGCTGTGGTACCCGACGATTGCTGGCAAAAAGGCTGCCG	2110
QY	101	GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu	120
Db	2111	CAGTCTGGTCAGGTGCGCATTAATTCCTATGTCCAGCGCGTACCCGTAACACGGCCAA	2170
QY	121	AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr	140
Db	2171	GCTTTTCGGCGCGGCGCTGGCACCTGACCTGTGCATTAACCGTACATCCACGACAGATGC	2236
QY	141	SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla	160
Db	2231	TCCAGTCCCGAATCCGTTATTATATCTCTTAAAAAATGGCGATTGGCACTGATACGCG	2290
QY	161	AsnValThrAspAlaIleLeuGlnLysArgAlaGlySerIleAlaAspPheThrGlyHis	180
Db	2291	AACGGAATGACGCGCATCTCTCAGCGAGGACAGAGAGGTCAATGTGCACTTTAACCGGGCAT	2350
QY	181	TyrglnThrAlaPheArgGluLeuGlnLysArgValIleAsnPheProGlnSerAsnLeuCys	200
Db	2351	CGGCAACGCGGCTTCGCGAATGTGAACGGGCTCTTAATTTTCCGCAATCAACTTGTGC	2410
QY	201	LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu	220
Db	2411	CTTAACGCGAAGAACAGACGAAAGCTGTTTCATTAAACGACGACATTACATCGGAATTC	2470
QY	221	LysValSerAlaAspCysValSerLeuThrGlyValAlaValSerLeuAlaSerMetLeuThr	240
Db	2471	AAGGTGAAGCGCGCAATGTCTCATTTAACCGGTGCGGTAAACCTCGCATCAATGCTGACG	2530
QY	241	GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr	260
Db	2531	GAGATATTTCTCTCTCAACAGACAGAGAAAGCCGAGCGGGGTGGGAAAGATCAAC	2590
QY	261	AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln	280
Db	2591	CATTACACACAGTGAACACCTGTGTAAGTTTGCATTAACGCGCAATTTTATTTGCTACA	2650
QY	281	ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla	300
Db	2651	CGCAGCGCAAGAGTGGCCGCGACGCGCGCACCCCGTTATTGAATTCATCAAGACAGCG	2710
QY	301	LeuThrProHisProProGluLysGlnAlaTyrglyValThrLeuProThrSerValLeu	320
Db	2711	TTGAGCGCCCAATCCACGCAAAAACAGCGATGGTGTGACATTACCACTTCAGTGTGCG	2770
QY	321	PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyValAlaLeuGluLeuAsnTrp	340
Db	2771	TTTATCGCGGACACGATCTAATGTGGCAATCTCGCGCGGCACTGGAAGTCAACTGG	2830
QY	341	ThiLeuProGluGlnProAspAsnThrProProGlyGlyGluLeuValPheGlnLysTrp	360

Db 2831 ACCCTTCCCGGTGACGCCGATTAACAGCCGCCGAGGTGTAACCTGGTTTGAACCTGG 2890

Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380

Db 2891 CGTCGGCTAAGCGATTAACAGCCAGTGATTCAGGTTTCGTGCTTCCAGACTTACAG 2950

Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400

Db 2951 CAGATCGGTGAATAAACGCCGCTGTCTAATAACCGCCGCCGAGAGGTGAACCTGACC 3010

Qy 401 LeuAlaGlyCysGlnGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420

Db 3011 CTGGCAGATGTGTAAGCAATGCGCAGGCGCATGTGTTCTGTGGCAGGTTTACGCCA 3070

Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432

Db 3071 ATCGTAATGAAGCAGCATACCGCTTGCAGTTTG 3106

RESULT 9
AX042373

LOCUS AX042373 6708 bp DNA linear PAT 23-NOV-2000

DEFINITION Sequence 2 from Patent WO0064247.

ACCESSION AX042373

VERSION AX042373.1 GI:11340991

KEYWORDS

SOURCE synthetic construct

ORGANISM synthetic construct

REFERENCE 1
Forseberg,C.W., Golovan,S. and Phillips,J.P.
Transgenic animals expressing salivary proteins
Patent: WO 0064247-A 2 02-NOV-2000;
UNIVERSITY OF GUELPH (CA)

FEATURES
source 1..6708
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="R15/APPA + intron plasmid with pBLCAT3 vector"

ORIGIN

Alignment Scores:
Pred. No.: 2,37e-195 Length: 6708
Score: 2182.00 Matches: 424
Percent Similarity: 98.1% Conservative: 0
Best Local Similarity: 98.1% Mismatches: 8
Query Match: 96.6% Indels: 0
DB: Gaps: 0

US-10-601-319-10 (1-432) x AX042373 (1-6708)

Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20

Db 1811 ATGAAGAGCATCTTAATCCCATTTTATCTCTTCATTCCTGTAACCCCGCATGTGA 1870

Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40

Db 1871 TTCGCTCAAGATGAGCGCGAGCTGGAAGCTGGAAGAGTGATGTCTCAGTCTCATGT 1930

Qy 41 ValArgAlaProThrLysValThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60

Db 1931 GTGCGTCTCCAAACGAGGCGACGCAATGATGCAAGATGCAACCCCAAGCATGGCA 1990

Qy 61 ThrTrpProValLysLeuGlyGluLeuThrProArgIleGlyGluLeuIleAlaTyrLeu 80

Db 1991 ACTGCGCCGGTAABAACGTGGGTGGCTGACACCGCGCGGTGGTGAAGTATTCGCTATTC 2050

Qy 81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100

Db 2051 GGACATTTACCAACGCGCAGGTCTGTGTAGCGCAACGGAATGCTGCGCAAAAAGGCGTCCG 2110

Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120

Db 2111 CAGTCTGTCAAGTCCGCTATTGCTGATGTGACAGCCGTAACCGCTAAACAGCCGAA 2170

Qy 121 AlaPheAlaIleGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140

Db 2171 GCCTTGGCCGGGGCTGGACCTGACTGTGCAATACCGTACATACCCAGGAGATAG 2230

Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160

Db 2231 TCCAGTCCCGATCCGTTATTTATCTCTPAAAAACGGGGGTGTGCGACATGGAATACGG 2290

Qy 161 AsnValThrAspAlaIleLeuGlnArgAlaGlyLysSerIleAlaAspPheThrGlyHis 180

Db 2291 AACGTACTGACGCGATCTCAGCAGGGGAGAGGGTCAATGTGACTTACCGGGCAT 2350

Qy 181 TyrGlnThrAlaPheArgLysLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200

Db 2351 CCGCAAAACGGCGTTTGGCGAATGGAAACGGGTGCTTAATTTCCGCAATCAAACTGTGC 2410

Qy 201 LeuLysArgGlnLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220

Db 2411 CTTAAACGTGAAGAAACAGGACGAAAGCTGTTCATTAAACGACAGCATTACATCGGAAC 2470

Qy 221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240

Db 2471 AAGGTAGCGCCGACATGTCTCATTAACCGGTGCGGTAGCTTCGCATCAATGCTGACG 2530

Qy 241 GlnIlePheLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrPylArgIleThr 260

Db 2531 GAGATATTTCTCTCGAACACAGCACAGGGAATCCCGAGCCGGGTGGGAAAGATCAC 2590

Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuGln 280

Db 2591 GATTCAACACAGATGGAACACTTGTGTAAGTTGCAATACGCCCAATTTATTTGCTACAA 2650

Qy 281 ArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300

Db 2651 CCGACGCCAGAGGTGGCCGCGCCGCGCACCCGTTATTAAGTTGATCAAGAGCGG 2710

Qy 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320

Db 2711 TTGACGCCCATCCACCGGAATAACAGGGGTATGGTGAATTCATTCACCTTCAGTGG 2770

Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyValAlaLeuGluLeuAsnTrp 340

Db 2771 TTTATCGCCGACAGATACCTAATCTGGCAATCTCGCGCGGCACTGGAGCTCAACTGG 2830

Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360

Db 2831 ACGCTTCCCGTCAACCGGATTAACAGCCGCCGAGGTGGTGAACCTGTGTAACGCTGG 2890

Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380

Db 2891 CGTCGGCTAAGCGATTAACAGCCAGTGAATTCAGGTTTCGTGCTTCCAGACTTACAG 2950

Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400

Db 2951 CAGATCGGTGAATAAACGCCGCTGTCTAATAACCGCCGCCGAGAGGTGAACCTGACC 3010

Qy 401 LeuAlaGlyCysGlnGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420

Db 3011 CTGGCAGATGTGTAAGCAATGCGCAGGCGCATGTGTTCTGTGGCAGGTTTACGCCA 3070

Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432

Db 3071 ATCGTAATGAAGCAGCATACCGCTTGCAGTTTG 3106

RESULT 10
AX042378

LOCUS AX042378 17732 bp DNA linear PAT 23-NOV-2000

DEFINITION Sequence 7 from Patent WO0064247.

ACCESSION AX042378

VERSION AX042378.1 GI:11340996

US-10-601-319-10 (1-432) x AX042372 (1-20623)

QY 1 MetLysAlaIleuLeuLeuProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 DB 12653 ATGAAGCCATCTTAATCCCATTTTATCTCTGATCCGTAACCCCGCAATCTGCA 12712
 QY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValAlaIleValSerArgHisGly 40
 DB 12713 TTGCGTCAGATGAGCGCGAGAGCTGGAAGCTGGAAAGTGTGTGATTTGTCACTGCTATGGT 12772
 QY 41 ValAlaGlnLeuProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 DB 12773 GTGCGGTCTCAACCAAGCGCAAGCAATGATGACAGATGTCACCCCAAGCGCATGGCA 12832
 QY 61 ThrTrpProValLysLeuGlnGlyLeuLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
 DB 12833 ACCGTGGCGGTAAACGTGGGTGGCTGACACCGCGGGGTGATGCTAATCGCTCATCTTC 12892
 QY 81 GlyHisTrpTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
 DB 12893 GGAATATACCAACGCCAGCGCTGTAGCCGACGATGTGCGCAAAAGGCTCCCG 12952
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGlnArgThrArgLysThrGlyGlu 120
 DB 12953 CAGTCTGTCAAGGTCCGATTAATGCTGATGTGACGAGGTACCCGTAACAGCGCA 13012
 QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 DB 13013 GCCTTCCGCGCGGGCTGGACCTGTGCATATACCGTACATCCAGGCGCATAG 13072
 QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAlaAla 160
 DB 13073 TCCAGTCCCGATCCGTAATTTAATCCTTAATAACCTGGCGTTGCCAATGATACGG 13132
 QY 161 AsnValThrAspAlaIleLeuGlnArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 DB 13133 AACGTGATCAGACCGCATCTCTCAGCAGGCGAGGGGTCAATTGCTACCTTAACGGGCA 13192
 QY 181 TyrGlnThrAlaPheArgGluLeuGlnLysValLeuAsnPheProGlnSerAsnLeuCys 200
 DB 13193 CGGCAAACCGCGTTCCGCAACCTGGAAACGGGTCTTAATTTCCGCAATCAAACTTGTC 13252
 QY 201 LeuLysArgGluLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 DB 13253 CTTAAACGTGAAGAACAGAACAACTGTCTTAACCAAGCATTAACATCGGAATCTC 13312
 QY 221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 DB 13313 AAGGTAGGCGCGCAATGCTCATTAACGGTGGGTAAAGCTCGCATCAATGCTGACG 13372
 QY 241 GluIlePheLeuLeuGlnGlnAlaGlnIleMetProGlnProGlyLysArgIleThr 260
 DB 13372 GAGATATTTCTCTCTGCACAAAGCACAGGATGCCGAGCCGGGGTGGGAAGATCAC 13432
 QY 261 AspSerHisGlnThrAsnThrLeuLeuSerLeuHisLeuAlaGlnPheAspLeuLeuGln 280
 DB 13432 GATTCACACAGAGGAAACCTTGCTAAGTTTGCAAAACGCAATTTATTATTGCTACAA 13492
 QY 281 ArgThrProGluValAlaArgSerAlaThrProLeuLeuAspLeuIleLysThrAla 300
 DB 13492 CGCAGCGCAAGAGTTGCCGCGCGCGCCGCTTATTAATTAATTAATTAATTAATTAATTA 13552
 QY 301 LeuThrProHisProProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
 DB 13552 TTGACGCCCCCATCAACCGCAAAACAGCGGTGCTGATGATCAATCCCACTCAAGTCTG 13612
 QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeuGluLeuAsnTrp 340
 DB 13612 TTTATCGCGGACACATATCTAATCTGGGAAATCTCGGGGGCGCATGAGAGCTCAACTGG 13672
 QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlyLeuValPheGlnArgTrp 360
 DB 13672 ACGCTTCCCGGTCAAGCGGATTAACAGCGCGCAAGGTGGAACTGGTGTTAACGCTGG 13732

QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 DB 13733 CGTGGCTAAGCGATTAACAGAGTGATTCAGGTTTCCTGGTCTTCCAGACTTTACAG 13792
 QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
 DB 13793 CAGATCGGTGATTAACCGCGCTGTCTAATTAATACGCCGCCGAGAGGTGAATCTGACC 13852
 QY 401 LeuAlaGlyCysGluGlnArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 DB 13853 CTGGCAGATGTGAAGAGCGGAATGCGCAGGCAATGTGTTCTGTGGCAGGTTTACGCA 13912
 QY 421 IleValAsnGlnValAlaArgIleProAlaCysSerLeu 432
 DB 13913 ATCGTAATGAAGACGATACCCGCTTGCAAGTTTG 13948

RESULT 12 U00096.10 WPCOMMENT

Sequence split into 47 fragments LOCUS U00096 Accession U00096

Fragment Name	Begin	End
U00096_00	1	110000
U00096_01	100001	210000
U00096_02	200001	310000
U00096_03	300001	410000
U00096_04	400001	510000
U00096_05	500001	610000
U00096_06	600001	710000
U00096_07	700001	810000
U00096_08	800001	910000
U00096_09	900001	1010000
U00096_10	1000001	1110000
U00096_11	1100001	1210000
U00096_12	1200001	1310000
U00096_13	1300001	1410000
U00096_14	1400001	1510000
U00096_15	1500001	1610000
U00096_16	1600001	1710000
U00096_17	1700001	1810000
U00096_18	1800001	1910000
U00096_19	1900001	2010000
U00096_20	2000001	2110000
U00096_21	2100001	2210000
U00096_22	2200001	2310000
U00096_23	2300001	2410000
U00096_24	2400001	2510000
U00096_25	2500001	2610000
U00096_26	2600001	2710000
U00096_27	2700001	2810000
U00096_28	2800001	2910000
U00096_29	2900001	3010000
U00096_30	3000001	3110000
U00096_31	3100001	3210000
U00096_32	3200001	3310000
U00096_33	3300001	3410000
U00096_34	3400001	3510000
U00096_35	3500001	3610000
U00096_36	3600001	3710000
U00096_37	3700001	3810000
U00096_38	3800001	3910000
U00096_39	3900001	4010000
U00096_40	4000001	4110000
U00096_41	4100001	4210000
U00096_42	4200001	4310000
U00096_43	4300001	4410000
U00096_44	4400001	4510000
U00096_45	4500001	4610000
U00096_46	4600001	4639675

Continuation (11 of 47) of U00096 from base 1000001 (U00096 Becherichia coli K-12 MG1655)

Alignment Scores: 1.42e-193 Length: 110000

Pred. No.:

Score: 2182.00 Matches: 424
 Percent Similarity: 98.1% Conservative: 0
 Best Local Similarity: 98.1% Mismatches: 8
 Query Match: 96.6% Indels: 0
 DB: 15 Gaps: 0

US-10-601-319-10 (1-432) x U00096_10 (1-110000)

```

Qy      1 Metlysalalieleulleperpholeuserleuileulleperoleuthrproginserla 20
Db      39840 ATGAAGGATCTTAATCCATTTTATCTCTTCGATTCGGTTAAACCCCAATCTGCA 39899
Qy      21 PhehlaaginsergluprogliuleuylseruuserValiilevaliserArgHisgly 40
Db      39900 TTCGCTCGAGTGAACCCGAGCTGAAGCTGGAAGCTGAGATTGTCAGTCGTCATGT 39959
Qy      41 ValArgalaProthrlyalaathrgluleuwerGlnaSerValThrProaspalaTrpPro 60
Db      39960 GTGCGTGCTCCAAACCAAGGCCACGCACTGATGCAAGATGTCAACCCAGACGCAATGGCCA 40019
Qy      61 ThrTrpProvallylsleuglygluleuthrProatrglygluleullealatytleu 80
Db      40020 ACCTGGCCGGTAAACCTGGGTGGCTGACACCGCGGGTGGAGCTTAATGCCCTATCTC 40079
Qy      81 GlyHisTyrTrpArgGlnArgleuValilaaspGlyleuLeuProlyeCysGlyCysPro 100
Db      40080 GGACATTTCCAACGCCGACGCTCTGTTACCGACGATTCGTGGCAAAAAGGGCTGCCCG 40139
Qy      101 GlnSerGlyGlnValalaileleilaasrValaaspGluatrgthrArglyserThrlyglu 120
Db      40140 CAGTCTGGTCAGTGTGGCATTTATGCTGATGTCGACGACGCTAACCCGTAAACAGCGCAA 40199
Qy      121 AlaPhealaalaglyleualaProasrCysalailethrValHisThrGlnalaasrThr 140
Db      40200 GCCTTCGCCCGCGGCTG3CACCTGACCTGTCAAATACCGTACATACCACGACATACG 40259
Qy      141 SerSerProasrProleuPheasrProleuylserThrlyValCysGlnleuAspAsnla 160
Db      40260 TCCAGTCCCGATCCGTTATTTTAACTCTTAAACCTGGCGTTTGCCAACTGGATACCGG 40319
Qy      161 AsnValThrAspalaileleugluarglaaglylserIlealaasrPheThrlyHis 180
Db      40320 AACGTAAGTGAAGGATCTCTGACGACGAGGAGGTCAATTGCTGACTTAAACGGGAT 40379
Qy      181 TyrGlnThrAlaPheatrglyleugluargvalleuAsrPheProGlnSerAsnleuys 200
Db      40380 CGGCAAAAGGGGTTTCGGAACCTGGAACGGGTGCTTAATTTTCCCAATCAAACTTGTGC 40439
Qy      201 LeuIysArgGluLysGlnAspGluSerCysSerleuThrGlnAlaLeuProSerGluLeu 220
Db      40440 CTTAAACCTGAGAAACAGACGAAAGCTGTTCAATTACGACGACATTACCATCGGAATC 40499
Qy      221 LysValSerAlaasrCysValaserleuThrlyAlaValSerleuAlaSerMetleuThr 240
Db      40500 AAGGTGACGCGCGCAATGTCTCATTTAACCGGTGCGTAAAGGCTGCATCAATGCTGACG 40559
Qy      241 GluIlePheleuLeuGlnGlnAlaGlnIlyMetProGluProGlyTyrGlyArgIleThr 260
Db      40560 GAGATATTTCTTCGCAACAGACAGGAAATGCGGAGCGGGGTGGGAAAGGATCAAC 40619
Qy      261 AspSerHisGlnTrpAsnThrleuLeuSerleuHisAsnAlaGlnPheAsrleuLeuGln 280
Db      40620 GATTCAKACCAAGTGAACACCTGTAGTTTGAATACGSCCAATTTTATTTGCTACAA 40679
Qy      281 ArgThrProGluValalaatGserArgAlaThrProleuLeuAsrleuIlelyserThxla 300
Db      40680 CGCACGSCAGAGGTTCGCGACGCGCGCACCCGTTATTAAGATTGATCAAGACAGACG 40739
Qy      301 LeuThrProHisArgProGlnIlyserIlyAlaTyrGlyValThrleuProThrSerValleu 320
Db      40740 TTGACGSCCATTCACCGCAAAACAGCGGTATGGTGTGACATTTACCACTTCAGTGTGG 40799
Qy      321 PheIleAlaGlyHisAsrThrAsrleuAlaAsrleuGlyAlaLeuGluIleuAsrThr 340

```

```

Db      40800 TTTATCCCGGACACGATTAATCTCGCAAAATCTCGCGCGGACCTGAGCTCAACTGG 40859
Qy      341 ThrleuProGlyGlnProasrPheThrProProGlygluleuValPheGluArgTrp 360
Db      40860 ACGCTCCCGGTGACCGCGATTAACACCGCCGACAGTGTGAATCGTGTGAACGCTGG 40919
Qy      361 ArgArgleuserAsrPaasrGlnTrpIleGlnValaserleuValPheGlnThrleuGln 380
Db      40920 CGTGGCTTAAGCGATTAACAGACGTGAATTCAGGTTTCGCTGTCCTTCAGACTTTACAG 40979
Qy      381 GlnMetArgAsrlyThrProleuSerleuAsnThrProProGlygluVallylsleuThr 400
Db      40980 CAGATGGGTGATTAACACCGCGCTGTCAATTAAATACCGCGCGGACGAGGTGAATCTGACC 41039
Qy      401 LeuAlaGlyCysGluGluArgAsnAlaGlnIlyMetCysSerleuAlaGlyPheThrGln 420
Db      41040 CTGGCAGAGATGTGAAGCGCAAAATGCGCAGGCAATGTGTGTTGGCAGGTTTTACGCAA 41099
Qy      421 IleValAsnGluAlaArgIleProAlaCysSerleu 432
Db      41100 ATCGTAATGAAGACCGCATACCGGCGTGCAGTTTG 41135

RESULT 13
AP009048_10
MPCOMMENT
Sequence split into 47 fragments LOCUS AP009048 Accession AP009048
Fragment Name Begin End
AP009048_00 1 110000
AP009048_01 100001 210000
AP009048_02 200001 310000
AP009048_03 300001 410000
AP009048_04 400001 510000
AP009048_05 500001 610000
AP009048_06 600001 710000
AP009048_07 700001 810000
AP009048_08 800001 910000
AP009048_09 900001 1010000
AP009048_10 1000001 1110000
AP009048_11 1100001 1210000
AP009048_12 1200001 1310000
AP009048_13 1300001 1410000
AP009048_14 1400001 1510000
AP009048_15 1500001 1610000
AP009048_16 1600001 1710000
AP009048_17 1700001 1810000
AP009048_18 1800001 1910000
AP009048_19 1900001 2010000
AP009048_20 2000001 2110000
AP009048_21 2100001 2210000
AP009048_22 2200001 2310000
AP009048_23 2300001 2410000
AP009048_24 2400001 2510000
AP009048_25 2500001 2610000
AP009048_26 2600001 2710000
AP009048_27 2700001 2810000
AP009048_28 2800001 2910000
AP009048_29 2900001 3010000
AP009048_30 3000001 3110000
AP009048_31 3100001 3210000
AP009048_32 3200001 3310000
AP009048_33 3300001 3410000
AP009048_34 3400001 3510000
AP009048_35 3500001 3610000
AP009048_36 3600001 3710000
AP009048_37 3700001 3810000
AP009048_38 3800001 3910000
AP009048_39 3900001 4010000
AP009048_40 4000001 4110000
AP009048_41 4100001 4210000
AP009048_42 4200001 4310000
AP009048_43 4300001 4410000
AP009048_44 4400001 4510000

```


CP000038_39 3900001 4010000
 CP000038_40 4000001 4110000
 CP000038_41 4100001 4210000
 CP000038_42 4200001 4310000
 CP000038_43 4300001 4410000
 CP000038_44 4400001 4510000
 CP000038_45 4500001 4610000
 CP000038_46 4600001 4710000
 CP000038_47 4700001 4810000
 CP000038_48 4800001 4825285
 Continuation (11 of 49) of CP000038 from base 1000001 (CP000038 Shigella sonnei Ss046, c

Alignment Scores:

Pred. No.: 1,42e-193 Length: 110000
 Score: 2182.00 Matches: 424
 Percent Similarity: 98.1% Conservatave: 0
 Best Local Similarity: 98.1% Mismatches: 8
 Query Match: 96.6% Indels: 0
 DB: 15 Gaps: 0

US-10-601-319-10 (1-432) x CP000038_10 (1-110000)

Qy 1 Metlysa1a1leu1leProPheLeuSerLeu1leProLeuThrProGlnSerAla 20
 Db 60260 ATGAAGGATCTTAATCCCATTTTATCTCTTGATTCGGTTAAACCCGCAATCTCA 60319
 Qy 21 PheAlaGlnSerGluProGluLeuLeuLysSerValIleValSerArgHisGly 40
 Db 60320 TTCCTCGAGTGTAGCCGAGACTGAAGTGAAGTGTGATGTCAGTCTCATAGT 60379
 Qy 41 ValArgAlaProThrLysAlaIleThrcGlnLeuMetGlnAspValIleThrProAspAlaTrpPro 60
 Db 60380 GTGCGTCTCCACCAAGCCAGCCAGCACTGATGACGAGTGTCAACCCGACGATGGCA 60439
 Qy 61 ThrTrpProValLysLeuGlyLysLeuThrProArgGlyGlyLysLeuIleAlaTyrLeu 80
 Db 60440 ACTGCGCGGTAAAACTGGTGTGGCTGACACCGCGGTGGTGAAGTAAATCGCTATCTC 60499
 Qy 81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLysLeuProLysCysGlyCysPro 100
 Db 60500 GACACTTACCAACGCGACGCTGTGTAGCCAGATGTGGCGAAAAAGGCGTGGCCG 60559
 Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgLysThrGlyGlu 120
 Db 60560 CAGTCTGCTCAGTCCGATTAATTCGTGATGCAAGCGATACCCCGTAACAGCGCA 60619
 Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 Db 60620 GCTTCGCGCGCGGCTGGCACTGATGCAATTAACCGTAACATACCCAGGCTGATACG 60679
 Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 Db 60680 TTCAGTCCCGAATCCGTATTTAATCTCTTAATAATGCGTTGCAACTGATTAACGCG 60739
 Qy 161 AsnValThrAspAlaIleLeuGlnLysArgAlaGlySerIleAlaAspPheThrGlyHis 180
 Db 60740 AACGATGATGACGCAATCTCTCAGCAGGCGCAGAGGCTCAATTCGATCTTAACGGGAT 60799
 Qy 181 TyrGlnThrAlaPheArgGluLeuGlnLysArgValLeuAsnPheProGlnSerAsnLeuCys 200
 Db 60800 CGGCAAAAGCGGCTTCGCAACTGGAACGGGCTGTAATTTTCCCAATCAACTTGTGC 60859
 Qy 201 LeuLysArgGluLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlnLeu 220
 Db 60860 CTTAAACGTGAAGAACAGACGAAAGCTGTTCATTAAAGCGAGGCACTTACATCGGAATCC 60919
 Qy 221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 Db 60920 AAGGTGACCGCGCAATGTCATTAACCGGTGGGTGAAGCTGATCAATGTCAGC 60979
 Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
 Db 60980 GAGATATTTCTCTTCAACAAGCAGAGGAATGCGGAGCGGCGGTGGGAAAGATCAC 61039

Qy 261 AsnSerHisGlnTrpAsnThrLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
 Db 61040 GATTCACACCAAGTGAACACTCTGTAAGTTTGATTAACCGCAATTTTATTTCTCTCA 61099
 Qy 281 ArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
 Db 61100 CGCACGCGCAGAGGTTCGCCGAGCGGCCACCCCGTTATTAATGATTCACAGACGCG 61159
 Qy 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
 Db 61160 TTGACGCCCATTCACCGCAAAAACAGCGCGTGTGTGACATTAAACCACTTCAGTGTG 61219
 Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTrp 340
 Db 61220 TTATTCGCGGACACGATTAATCTGCAATATTCGCGCGGCGCATCTGAGCTCAACTGG 61279
 Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 Db 61280 ACGCTTCCCGGTCAACCGGATTAACGCGCGCAGGTGTGAACTGTGTTGAACGCTGG 61339
 Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 Db 61340 CGTGGCTTAAGCGATTAACAGCCAGTGTATTCAGGTTCGCTTCACAGACTTTACAG 61399
 Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 Db 61400 CAGATGCGTGTAAAGCCGCTGTCAATTAAATACGCCCGCGAGAGGTCAAACTGACC 61459
 Qy 401 LeuAlaGlyCysGlnGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 Db 61460 CTGGCAGATGTGAAGCGCAATATGCCAGGCGATGTGTGTTGCGAGGTTTACGCA 61519
 Qy 421 IleValAsnGlnAlaArgGlyLeuProAlaCysSerLeu 432
 Db 61520 ATCGTAATGAAGCACCATATCCGCGTGCAGTTTG 61555

RESULT 15

LOCUS AR636188 1901 bp DNA linear PAT 20-APR-2005
 DEFINITION Sequence 5 from patent US 6855365.
 ACCESSION AR636188
 VERSION AR636188.1 GI:62768871
 KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE Unclassified.

1 (bases 1 to 1901)
 Short, J.M., Kretz, K.A., Gray, K.A., Barton, N.R., Garrett, J.B.,

O'Donoghue, E. and Mathur, B.J.

Recombinant bacterial phycases and uses thereof

Patent: US 6855365-A 5 15-FEB-2005;

Diversa Corporation; San Diego, CA

Location/Qualifiers

1..1901

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 4.65e-196 Length: 1901
 Score: 2181.00 Matches: 424
 Percent Similarity: 98.1% Conservatave: 0
 Best Local Similarity: 98.1% Mismatches: 8
 Query Match: 96.6% Indels: 0
 DB: 2 Gaps: 0

US-10-601-319-10 (1-432) x AR636188 (1-1901)

Qy 1 Metlysa1a1leu1leProPheLeuSerLeu1leProLeuThrProGlnSerAla 20
 Db 188 ATGAAGGATCTTAATCCCATTTTATCTCTTGATTCGGTTAAACCCGCAATCTCA 247

QY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValAlaIleValSerArgHisGly 40
 DB 248 TTCGTCAGAGTGAAGCCGAGAGCTGAAGCTGGAAGAGTGTGTGATTTGTCATGCTCATGCT 307
 QY 41 ValAlaGlnProThrThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 DB 308 GTGCGGTCTCCAAACCAAGGCCACGCAATGATGACAGATGTCACACCCCAAGCGCATGGCCA 367
 QY 61 ThrTrpProValLysLeuGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
 DB 368 ACCCTGGCCGTAACCTGGGTGACCTGACCAACCGCGAGGTGTAGCTGAATTCGCTCATCTTC 427
 QY 81 GlyHisTrpTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
 DB 428 GGAACATTACCAACGCCAGGCTCTGGTACCGAGGATTCGTGGCGAAGAAAGGCTGCCCCG 487
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 DB 488 CAGTCTGTCAAGTCCGATTTATTCGATGTCGACGAGGTACCCGTAACCCGTAACAGCGCAA 547
 QY 121 AlaPheAlaAlaGlyLeuLeuLysProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 DB 548 GCTTTCGCGCGCGGTGGCACTGTGCAATACCGTACATACCCAGCGCAGATACG 607
 QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 DB 608 TTCAGTCCCGATCCGTTATTTAATCTCTTAACAACTGGCGTTTGGCACTGGATACGCG 667
 QY 161 AsnValThrAspAlaIleLeuGlnArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 DB 668 AACGTGACTGACGCGATCTCTCAGCAGGCGCAGAGGGGTCAATTGCTGACTTTACCGGGCAT 727
 QY 181 TyrGlnThrAlaPheArgGluLeuGlnLysValLeuAsnPheProGlnSerAsnLeuCys 200
 DB 728 CGGCAAACGCGCTTTCGCACTGGAAACGCGGTCTTAATTTCCGCAATCAAACTTGTGC 787
 QY 201 LeuLysArgGluLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 DB 788 CTTAAACGTGAAGAACAGACGAAAGCTTTCAATTACCGAAGCATTACCATCGAATCTC 847
 QY 221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 DB 848 AAGGTAGCGCCGACAAATGCTCATTTAACCGGTGCGTAAGCTGCAATGCTGACG 907
 QY 241 GlnIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
 DB 908 GAGATATTTCTCTGTGCAACAGCACGAGGAATGCCGAGCCGCGGTGGGAAAGATCACCC 967
 QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
 DB 968 GATTCAACACGATGAACACCTTGTCTAAGTTGCAATACCGGCAATTTATTTGCTACAA 1027
 QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
 DB 1028 CGCAGCGCAGAGGTGGCCGCGCGCGCCGCAACCCGTTATTTAGATTGATCAAGACAGCG 1087
 QY 301 LeuThrProHisProProGlnLysGlnAlaIleGlyValThrLeuProThrSerValLeu 320
 DB 1088 TTGAGCGCCCATCCACCGCAAAACAGCGGTATGTGTGACATTACCCCACTTCAGTGTG 1147
 QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTrp 340
 DB 1148 TTTATCGCCGACAGATCTAATCTGCGAATCTCGGGGCGCACTGAGGCTCAACTGG 1207
 QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGlnArgTrp 360
 DB 1208 ACGCTTCCCGGTCAAGCCGGAATACAGCGCGCAGGAGGTGAACCTGGTTTGAACGCTGG 1267
 QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 DB 1268 CGTGGCTTAAGGATTAACGACGATGATTCAGGTTTCGTGCTTCACAGACTTTACAG 1327
 QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400

DB 1328 CAGATCCGATGAATAACGCCGCTGTCTAATATACGCCGCCGAGAGGTGAATACTGACC 1387
 QY 401 LeuAlaGlyCysGluGluLysArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 DB 1388 CTGGCAGATGTGAAGAGCGAAATGCGCAGGCGCATGTGTTGTTGCAAGTTTACGCAA 1447
 QY 421 IleValAsnGluLysArgIleProAlaCysSerLeu 432
 DB 1448 ATCGTAATGAAGCAGCATACCGCGGTGCAAGTTTG 1483

Search completed: June 14, 2006, 13:55:38
 Job time : 5313.22 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 14, 2006, 10:38:44 ; Search time 552.385 Seconds
(without alignment)
8179.104 Million cell updates/sec

Title: US-10-601-319-10

Perfect score: 2258

Sequence: 1 MALLIFPLSLILPLTPQSA.....CSLAGPTQVNEARIRACSL 432

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10469840

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Maximum Match 0%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-O=/abs/ABSSMB.spool/US10601319/runat.12062006.180047.27417/app.query.fasta_1
-DB=N Geneseg -OPT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bites -START=1 -END=-1 -MATRIX=blomsun62 -TRANS=human40.cdi -1157=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs806h
-USER=US10601319 @CCN 1_1 749 @runat.12062006.180047.27417 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV TIMEOUT=120
-MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DLEOP=6 -DLEXT=7

Database :

N_Geneseg_8: *
1: geneseqn1980s: *
2: geneseqn1990s: *
3: geneseqn2000s: *
4: geneseqn2001as: *
5: geneseqn2001bs: *
6: geneseqn2002as: *
7: geneseqn2002bs: *
8: geneseqn2003as: *
9: geneseqn2003bs: *
10: geneseqn2003cs: *
11: geneseqn2003ds: *
12: geneseqn2004as: *
13: geneseqn2004bs: *
14: geneseqn2005s: *
15: geneseqn2006s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	2258	100.0	1296	14	AED50817 Escherich
2	2258	100.0	1308	10	ADC87742 DNA encod
3	2258	100.0	1308	12	AD050299 Escherich

SUMMARIES

4	2258	100.0	1308	15	AEE75420 E. coli B
5	2258	100.0	1901	15	AEE75422 E. coli B
6	2195	97.2	1901	12	AD050296 Escherich
7	2190	97.0	1901	12	AD050295 Escherich
8	2190	97.0	1901	15	AEE75416 E. coli P
9	2182	96.6	1299	3	ABK12514 DNA encod
10	2182	96.6	1299	8	ACA19297 Prokaryot
11	2182	96.6	1299	12	ADL16137 Escherich
12	2182	96.6	1901	6	ADL25463 Escherich
13	2182	96.6	1901	8	ADA19449 E. coli K
14	2182	96.6	1901	10	ADC87744 DNA encod
15	2182	96.6	1901	12	AD050297 Escherich
16	2182	96.6	1901	14	AED50819 Escherich
17	2182	96.6	1901	15	AEE75418 E. coli B
18	2182	96.6	3470	3	ACG68298 R15/APPA
19	2182	96.6	4060	3	ACG68296 R15/APPA
20	2182	96.6	6116	3	ACG68297 R15/APPA
21	2182	96.6	6708	3	ACG68295 R15/APPA
22	2182	96.6	17732	3	ACG68300 Lama2/App
23	2182	96.6	20623	3	ACG68294 Lama2/App
24	2177	96.4	1296	12	ADL16139 Shigella
25	2175	96.3	1901	15	AEE75417 E. coli P
26	2172	96.2	1323	12	AD050303 Escherich
27	2171	96.1	1323	5	ACG88885 Escherich
28	2171	96.1	1323	6	ADJ36473 Escherich
29	2171	96.1	1323	6	ADJ25460 Escherich
30	2171	96.1	1323	8	ADA19445 E. coli B
31	2171	96.1	1323	12	AD050291 Escherich
32	2171	96.1	1323	15	AEE75412 E. coli B
33	2169	96.1	1323	12	AD050301 Kangaroo
34	2168.5	96.0	1296	14	AED46567 Nov9x phy
35	2168.5	96.0	1313	9	ADJ57148 pNOV4054
36	2168.5	96.0	1314	14	AED46569 Nov9x phy
37	2168.5	96.0	1331	9	ADJ57149 pNOV4058
38	2168.5	96.0	11357	14	AEC39118 pNOV4061
39	2165	95.9	1315	10	ADL02197 DNA encod
40	2164	95.8	1486	5	ADJ06831 E. coli a
41	2161	95.7	1256	9	ADJ57147 Maize-opt
42	2157	95.5	1281	9	ACF35789 Nov9x phy
43	2152	95.3	1901	8	ADA19452 E. coli K
44	2152	95.3	5421	3	ACG68299 SV40/APPA
45	2150	95.2	1296	12	ADL16138 Shigella

ALIGNMENTS

RESULT 1	
ID	AED50817 standard; DNA; 1296 BP.
AC	AED50817;
XX	
XX	
DT	29-DEC-2005 (first entry)
XX	
DE	Escherichia coli phytase appA mutant DNA SEQ ID NO: 1 #1.
KW	Transgenic animal; gene therapy; feedstuff; food; thermostable; mutant;
KW	ds; coding sequence; 3-phytase; recombinant DNA; appA gene.
XX	
OS	Escherichia coli; strain K12.
XX	
XX	Synthetic.
FT	Key
FT	Location/Qualifiers
FT	1..1296
FT	/tag= a
FT	/EC number= "3.1.3.8"
FT	/partial
FT	/product= "Escherichia coli phytase mutant protein"
FT	/note= "No stop codon"
PN	US2005246780-A1.
XX	
PD	03-NOV-2005.

XX 11-FEB-2005; 2005US-00056354.
PF 13-AUG-1997; 97US-00910796.
XX 01-MAR-1999; 99US-00259214.
PR 13-APR-1999; 99US-00291931.
PR 25-MAY-1999; 99US-00318528.
PR 25-MAY-2000; 2000US-00580515.
PR 24-MAY-2001; 2001US-00865379.
PR 24-MAY-2002; 2002US-00156660.
XX (SHOR/) SHORT J M.
PA (KRETZ/) KRETZ K.
PA (GRAY/) GRAY K A.
PA (BART/) BARTON N R.
PA (GAR/) GARRETT J B.
PA (ODON/) ODOONOHUE E.
PA (MATH/) MATHUR E J.
XX
PI Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, Odonoghue E,
PI Mathur EJ;
XX WPI: 2005-733851/75.
DR P-PSDB; AED50818.
XX
XX New isolated or recombinant nucleic acid encoding phytase enzymes, useful
PT as foodstuff, and for oil degumming, producing an animal feed, delivering
PT a phytase enzyme supplement to an animal.
XX
PS Claim 1; SEQ ID NO 1; 104pp; English.
XX
XX The present invention relates to a recombinant phytase (EC 3.1.2.8)
CC protein and its encoding also nucleic acid molecule. Phytases of the
CC invention is a thermostable protein. They are useful in foodstuffs, for
CC oil degumming, producing an animal feed, delivering a phytase enzyme
CC supplement to an animal, increasing the resistance of a phytase
CC polypeptide to enzymatic inactivation in a digestive system of an animal
CC and in processing of corn and sorghum kernels. The invention is also
CC useful in gene therapy and in production of transgenic animals. The
CC present sequence is a Escherichia coli phytase appa mutant DNA. Note: The
CC present sequence is the SEQ ID NO: 1 which is given in the sequence
CC listing. This sequence differs from the SEQ ID NO: 1 shown on page 19 in
CC disclosure of the specification (see AED50824).
XX
XX
SQ Sequence 1296 BP; 311 A; 342 C; 356 G; 287 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 3.43e-202 Length: 1296
Score: 2258.00 Matches: 432
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 14 Gaps: 0
US-10-601-319-10 (1-432) x AED50817 (1-1296)
QY 1 MetLvsAlaIleIleuIleProPhelenuSerLeuIleuIleProLeuThrProGlnSerAla 20
DB 1 ATGAAGGAGATCTTAATCCCATTTTATCTCTTGATTCGATVACCCCGAATCTCA 60
QY 21 PheIaGlnSerGluProGluLeuIleuIleuSerValIleIleValSerArgHisGly 40
DB 61 TTCGCTCAGAGTGAACCCGAGCTGAAGCTGAAAGTGTGATGATGTCAGTGTCAATG 120
QY 41 ValAlaGAlaProThrIleuAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
DB 121 GTGCTGTCTCCAAACAGGCCACGCACTGATGCGAGTGTCAACCCCGACATGGCCA 180
QY 61 ThrTrpProValIleuLeuGlyGluLeuThrProArgGlyGlyGluLeuIleuIleuIleu 80
DB 181 ACCTGGCCGGTAAACCTGGGTGAGCTGACACCGCCGGGTGGTGAAGCTTAATCCCTATCTC 240
QY 81 GlyHisTyrtTrpArgGlnArgLeuValIleAspGlyLeuLeuProLysCysGlyCysPro 100

DB 241 GGACATTACTGCGCGTACGCGTCTGGTAGCCACGCAATGGCTGCTTAATGTGGCTGCCG 300
QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLeuThrGlyGlu 120
DB 301 CAGTCTGGTCAAGTCTGCGCATTTATGCTGATGTCGACGAGCGCTACCCGTAACAGGCCAA 360
QY 121 AlaPheAlaIleuAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
DB 361 GCTTGGCCGCGGCTGGCACCTGACTGTGCATTAACGTAACATTAATCCAGGCAGATACG 420
QY 141 SerSerProAspProLeuPheAsnProLeuIleuThrGlyValCysGlnLeuAspAsnAla 160
DB 421 TCCAGTCCCGATCCGTTATTTATCTTAATCTTAACAACTGGCGCTTGGCACTGATACCGC 480
QY 161 AsnValThrAspAlaIleLeuGlnArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
DB 481 AACGTGACTGACGCGATCTCGAGAGGCGCAGAGGGTCAATGCTGACTTAACCGGCGAT 540
QY 181 TyrGlnThrAlaPheArgGluLeuGlnArgValLeuAsnProGlnSerAsnLeuGly 200
DB 541 TATCAACGCGCTTTCGCAACTGGAACGGGTGCTTAATTTCCGCAATCAACTGTGCG 600
QY 201 LeuIleArgGlyIleuIleuAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
DB 601 CTTAAAGCTGAAGAAACAGAGCAAGAAAGCTTCAATTAAAGCAGGCAATTACATCGGAATC 660
QY 221 LysValSerAlaAspCysValSerLeuThrGlyValAlaSerLeuAlaSerMetLeuThr 240
DB 661 AAGGTAGCGCGCGACTGTGTCTCATTAACCGGTGCGGTAACCTTCGCAATCATGCTGACG 720
QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
DB 721 GAGATATTTCTCTCGCAACAGCACAGGAAATGCCAGCGCGGGTGGGAAAGATCAC 780
QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
DB 781 GATTACACACGAGTGAACACCTTGCTAAGTTTGATTAACGCGCAATTTGCTGACAA 840
QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleuIleThrAla 300
DB 841 CGCAGCGCAAGAGTGGCCCGCGAGCGCGCCACCCCGTATTAAGATTGATCAAGACAGCG 900
QY 301 LeuThrProHisProProGlnIleuGlnAlaIleArgIleValThrLeuProThrSerValLeu 320
DB 901 TTGAAGCCCATTCACCGCCAAACAGCGCGTATGCTGTGACATTAACCATTTCACTGCTG 960
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
DB 961 TTTATCGCCGAGACAGATACTAATCTGGCAATCTCGCGCGCGCACTGGAGCTCAACTGG 1020
QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGlnArgTrp 360
DB 1021 ACGCTTCCCGGTACGCGCGATTAACACGCGCGCAGGTGTGAATGCTTTGAACGCTGG 1080
QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
DB 1081 GTCGCGCTTAAGCGATTAACAGCCAGTGAATTCAGGTTTCGCTGTCTTCCAGACTTTACAG 1140
QY 381 GlnMetArgAspLysTrpProLeuSerLeuAsnThrProProGlyGlyValIleLeuThr 400
DB 1141 CAGATGGGTGATTAACAGCGCGCTGTCAATTAAATAGCGCGCGGAGAGGTGAACCTGACC 1200
QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
DB 1201 CTGGCAGGATGTGAAGCGAAATGCGCAGGGCAATGTGTGTTGGCAGGTTTACGCA 1260
QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
DB 1261 ATCGTAATGAAGCAGCATACCGGCTGCAAGTTTG 1296
RESULT 2
ADCB7742

ID ADC87742 standard; DNA; 1308 BP.
AC ADC87742;
XX
XX
DT 01-JAN-2004 (first entry)
XX
DE DNA encoding the modified *Escherichia coli* phytase.
XX
XX phytase; food supplement; enzyme delivery matrix; soybean meal;
KW chemo-tolerance; thermostability; kernel; phytase;
KW myo-inositol-hexaphosphate; inositol; inorganic phosphate;
KW thermotolerant; feed value; digestion; gene; ds.
XX
OS Synthetic.
OS *Escherichia coli*.
XX
XX
FH Key Location/Qualifiers
FT 1..1308
FT /tag= a
FT /product= "Modified phytase"
PN US2003103958-A1.
XX
XX
PD 05-JUN-2003.
XX
XX
PF 24-MAY-2002; 2002US-00156660.
XX
PR 13-AUG-1997; 97US-00910798.
PR 13-MAR-1999; 99US-00259214.
PR 13-APR-1999; 99US-00291931.
PR 25-MAY-1999; 99US-00318528.
PR 25-MAY-2000; 2000US-00580515.
PR 24-MAY-2001; 2001US-00866379.
XX
XX (DIVE-) DIVERSA CORP.
PI Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI Mathur EJ;
XX
XX WPI. 2003-787039/74.
DR P-ESDB; ADC87743.
XX
XX
PT New nucleic acid encoding a polypeptide having phytase activity, useful
PT in improving the feeding value of phytate rich ingredients or as an aid
PT in phytate digestion.
XX
XX
PS Claim 9; SEQ ID NO 1; 113pp; English.
XX
XX
XX The invention discloses a new isolated or recombinant nucleic acid which
XX encodes a polypeptide having a phytase activity. Also claimed is a
XX nucleic acid probe, an amplification primer sequence pair, an expression
XX cassette comprising the nucleic acid, a vector comprising the nucleic
XX acid, a transgenic non-human animal or plant, or its seed, comprising the
XX nucleic acid, an antisense oligonucleotide, inhibiting the translation of
XX a phytase message in a cell, a heterodimer comprising the polypeptide and
XX a second domain, an array comprising immobilised polypeptide or nucleic
XX acid, a hybridoma comprising an antibody that specifically binds to the
XX polypeptide, a food supplement for an animal, an edible enzyme delivery
XX matrix, an edible pellet comprising a granule edible carrier and the
XX polypeptide, a feed composition, a soybean meal, isolating or identifying
XX the polypeptide, making an anti-phytase antibody, producing a recombinant
XX polypeptide, determining whether a compound binds to the polypeptide,
XX identifying a modulator, whole cell engineering of new or modified
XX phenotypes by using real-time metabolic flux analysis, increasing
XX thermo-tolerance or thermostability of the phytase polypeptide, increasing
XX the resistance of the phytase polypeptide to enzymatic inactivation in a
XX digestive system and processing of corn and sorghum kernels. The phytase
XX activity comprises catalysis of phytate (myo-inositol-hexaphosphate) to
XX inositol and inorganic phosphate and is thermotolerant. The nucleic acid
XX is useful in improving the feeding value of phytate rich ingredients or
XX as an aid in phytate digestion. The sequence presented is the DNA
XX encoding the modified *Escherichia coli* phytase.

Sequence 1308 BP; 315 A; 345 C; 357 G; 291 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 3,47e-202 Length: 1308
Score: 2258.00 Matches: 432
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 10
US-10-601-319-10 (1-432) x ADC87742 (1-1308)
QY 1 MetLySAIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAAGCATCTTAATCCATTATCTCTTGATTCGTAAACCCGCAATCTGCA 60
QY 21 PhealaginSerGluProGluLeuLeuSerValIleValSerArgHisgly 40
Db 61 TTCGCTCAGAGTGAACCCGAGCTGAAGCTGAAAGTGTGTATGTGCTGCTATGCT 120
QY 41 ValArgAlaProThrLySAIleThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 121 GTGCGTGTCTCCAAACCAAGGCCACGCACTGATGACGATGTCACCCGACCATGGCCA 180
QY 61 ThrTrpProValIleLeuGluGluLeuThrProArgIleGluLeuIleAlaTyrlleu 80
Db 181 ACCTGGCGGTAAACCTGGGTGAGTGAACACCGCGGTGGAGCTAATCGCTAATCTTC 240
QY 81 GlyHisTyrlleuProGluLeuValAlaAspGlyLeuLeuProLySCysGlyCysPro 100
Db 241 GGACATTAATCGGCGTCAAGCTGTGTACCGACGGAATGTGCTCCCTTAATGTGGTCCCG 300
QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAlaAspGluArgThrArgLyThrGlyGlu 120
Db 301 CAGTCTGTGTGAGTGTGCGCATTAATGCTGATGTCAGACGCGTAAACCGCGCA 360
QY 121 AlaPheAlaIleGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCTTCGGCGCGGCTGTGACCTGACCTGTGAATACCGTACATTAACCCAGCAATACG 420
QY 141 SerSerProAspProLeuPheAsnProLeuLyThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCCAGTCCCGAATCCGTTATTTAATCTCTTAATAACGTGCGCTTGCCAACTGATACCGC 480
QY 161 AsnValThrAspAlaIleLeuGluArgAlaGlySerIleAlaAspPheThrIleHis 180
Db 481 AACGTGACTGACGCGATCTCGAAGGGCAGAGGGTCAATGTCGATCTTAACCGGGCAT 540
QY 181 TyrGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 541 TATCAACGGGCTTCCGGAACCTGAAACGGGTGCTTAATTTTCCGCAATCAAACTTGGC 600
QY 201 LeuLySArGluLySGlnAspGlySerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAAACCTGAGAAACAGAGAAAGCTGTCTTAACGACGACATTAACCTGGAATC 660
QY 221 LySAIleSerAlaAspCysValSerLeuThrIleValAlaSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGAGCGCGCATGTGTCTCATTAACCGTGGTGAACCTCGCATCAATGCTGACG 720
QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Db 721 GAGGATTTCTCTGCAACCAAGCACAGGAATGCGGAGCGGGGTGGGAAAGATCACCC 780
QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuGln 280
Db 781 GATTCAACCAAGTGAACCACTTCTTAAGTTTGATTAACCGGCAATTTGATTTCTCA 840
QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLyThrAla 300
Db 841 CGCAGCGCAGAGTGTGGCCGCGACGCGCGCCACCCGTTAATTAAGATTGATCAAGACGCG 900
QY 301 LeuThrProHisProProGlnIleGlnAlaTyrlleValThrLeuProThrSerValleu 320

Db	901	TTGAGCCCATCCACCGCAAAAACAGCGGATGGTGTGACATTACCACTTCAGTGGCTG	960
Qy	321	PheIlealagIyhiSaSPThrAsnleuAlaSnleuGIyGIyAlaIleuGIuIeuaSnTrp	340
Db	961	TTTATCGCGGACACGATACTATATCTGGCAATCTCGCGGGGCACTGAGCTCACTGG	1020
Qy	341	ThrlieuP-roGIyGIuInProAspaSnThrProP-roGIyGIyGIuIeuValPheGIuIaGrTrp	360
Db	1021	ACGCTTCCCGGTCACCGCGATPACACGCGCGCAGGTGGTGAATCTGTGATTTGAACGCTGG	1080
Qy	361	ArgArgIeuSerAspaSnSerGIuInTrpIleGIuValSerIeuValPheGIuInThrIeuGIuIn	380
Db	1081	CGTCGGCTAAGAGATPACAGCCAGTGGATTTCAGGTTTCGCTGGTCTTCACAGCTTTACAG	1140
Qy	381	GlnMetAlaGaaPlyrTrpThrProIeuSerIeuSnThrProP-roGIyGIuIuValIlyIeuThr	400
Db	1141	CAGATGCGGTGATAAAACCGCTGTGCATTAAATAGCCGCCCGGAGAGGTGAATGACCC	1200
Qy	401	IeuAlaGIyCySGluGIuArgGaaAlaGIuInGIyMetCySserIeuAlaGIyPheTrGIuIn	420
Db	1201	CTGGCAGAGATGTGAAGACGAGAAATCGCAGGGCAGTGTCTTCCTTGCGAGGTTTTACGCA	1260
Qy	421	IleValaSnGIuAlaArgIleProAlaCySserIeu	432
Db	1261	ATCGTGAAATGAAGCAGCATACCGCGCTGCAGTTGG	1296

RESULT 3	
AD050299	ID
AD050299	standard; DNA; 1308 BP.
XX	AC
XX	AD050299;
XX	DT
XX	29-JUL-2004 (first entry)
XX	DE
XX	Escherichia coli K12 appa phytase mutant 819PH59 DNA.
XX	DB
KW	Appa phytase; bacteria; thermal tolerance; protease stability; foodstuff;
KW	phytate; animal feed; fish feed; dough; baking; gene; mutant; ds.
XX	
OS	Escherichia coli; K12.
OS	Synthetic.
XX	
PH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..1296
FT	/*tag= a
FT	/product= "Mutant phytase"
FT	/note= "Does not include stop codon"
FT	/partial
FT	replace(202,T)
FT	/*tag= b
FT	/note= "Corresponds to position 389 in the wild type
FT	sequence"
FT	replace(203,G)
FT	/*tag= c
FT	/note= "Corresponds to position 390 in the wild type
FT	sequence"
FT	replace(250,C)
FT	/*tag= d
FT	/note= "Corresponds to position 437 in the wild type
FT	sequence"
FT	replace(251,A)
FT	/*tag= e
FT	/note= "Corresponds to position 438 in the wild type
FT	sequence"
FT	replace(252,A)
FT	/*tag= f
FT	/note= "Corresponds to position 439 in the wild type
FT	sequence"
FT	replace(255,C)
FT	/*tag= g
FT	/note= "Corresponds to position 442 in the wild type
FT	sequence"

PT	mutation	replace (293,G) /*tag= h /note= "Corresponds to position 470 in the wild type sequence" mutation replace (285,G) /*tag= i /note= "Corresponds to position 472 in the wild type sequence" mutation replace (289,G) /*tag= j /note= "Corresponds to position 476 in the wild type sequence" mutation replace (290,A) /*tag= k /note= "Corresponds to position 477 in the wild type sequence" mutation replace (291,G) /*tag= l /note= "Corresponds to position 478 in the wild type sequence" mutation replace (502,A) /*tag= m /note= "Corresponds to position 689 in the wild type sequence" mutation replace (503,G) /*tag= n /note= "Corresponds to position 690 in the wild type sequence" mutation replace (504,C) /*tag= o /note= "Corresponds to position 691 in the wild type sequence" mutation replace (541,C) /*tag= p /note= "Corresponds to position 728 in the wild type sequence" mutation replace (542,G) /*tag= q /note= "Corresponds to position 729 in the wild type sequence" mutation replace (543,G) /*tag= r /note= "Corresponds to position 730 in the wild type sequence" mutation replace (676,A) /*tag= s /note= "Corresponds to position 863 in the wild type sequence" mutation replace (677,A) /*tag= t /note= "Corresponds to position 864 in the wild type sequence" mutation replace (829,T) /*tag= u /note= "Corresponds to position 1016 in the wild type sequence"
PA	(GRAY//) GRAY K A.	
PA	(SHOR//) SHORT J M.	
PA	(KRETZ//) KRETZ K.	
PA	(GRAY//) GRAY K A.	

PA (SHOR/) SHORT J M.
PA (KRET/) KRETZ K.
PA (GRAY/) GRAY K A.

PA (BART/) BARTON N R.
PA (GARR/) GARRETT J B.
PA (ODON/) O'DONOGHUE E.
PA (MATH/) MATHER E J.
PI Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI Mather EJ;
DR MPI: 2004-374952/35.
DR P-PSDB; ADOS0300.
PT Producing phytase, involves providing nucleic acid derived from bacteria
PT encoding polypeptide having phytase activity, and expressing nucleic acid
PT in yeast.
XX
XX
PS Claim 3; SEQ ID NO 9, 74pp; English.
CC The invention relates to a method for producing phytase that involves
CC providing a nucleic acid encoding phytase derived from a bacteria, and
CC expressing the nucleic acid in a yeast under conditions that allow
CC expression of the enzyme in the yeast. The invention also relates to
CC modified phytase enzyme which has improved thermal tolerance and protease
CC stability at low pH. The phytase enzyme can be used in foodstuffs to
CC improve the feeding value of phytate rich ingredients, and in diet of
CC numerous animals including mammals, fowls and fishes, commercially
CC significant mammals such as pigs, goats, laboratory rodents, commercially
CC significant avian species such as chicken, ducks, doves, parrot, etc.,
CC commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC etc., in dough making and baking, in dietary aids for animals. The method
CC provides easy manufacture of the active ingredient loaded biocompatible
CC composition, higher yields and loading efficiency. The phytase
CC incorporated in to the dietary aid is safe for animals. The present
CC sequence is Escherichia coli K12 apba phytase mutant DNA.

Sequence 1308 BP; 315 A; 345 C; 357 G; 291 T; 0 U; 0 Other;

Alignment Scores:

Align. No.: 3, 47e-202 Length: 1308
Score: 2258.00 Matches: 432
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 12 Gaps: 0

US-10-601-319-10 (1-432) x ADOS0299 (1-1308)

QY 1 MetLysAlaIleLeuIleProHeuSerLeuLeuProLeuThrProGlnSerAla 20
DB 1 ATGAAGCGATCTTAATCCATTTTATCTTCGATTCCGTTAACCCCGCAATCTGA 60
QY 21 PheAlaGlnSerGluProGluLeuLeuGlnSerValIleValSerArgHisGly 40
DB 61 TTCGTCGAGTACCGCGAGCTGAAGCTGAAGTGTGATGTCAGTGCATGGT 120
QY 41 ValArgAlaProThrThyValIleThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
DB 121 GTGCGTCTCCCAACCAAGCCACGCACTGATGCGAGATGCAACCCCAACCCATGGCCA 180
QY 61 ThrTrpProValIleLeuGlnIleuLeuThrProArgGlyGlyGlnLeuIleAlaTyLeu 80
DB 181 ACCTGGCCGGTAAACTGGGTGAGCTGACACCGCGGGTGGTGAAGTATGCCCTATCTC 240
QY 81 GlyHisGlyTrpTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
DB 241 GGACATTAATGCGCTGACGCGTCTGGTAGCGACGATTCCTCAATATGTCGTCGCCG 300
QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValaAspGluArgThrArgGlyThrGlyGln 120
DB 301 CAGTCTGGTCAAGCTGCGCATTTTGTCTGATGTCGACGACCGTAAACCAAGCCGAA 360
QY 121 AlaPheAlaIleGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
DB 361 GCCTTCGCGCGCGGCTGGCACTGACTGTGCAATTAACCGTATCCAGGCAATACG 420

QY 141 SerSerProAspProLeuPheAsnProLeuLysIleValIleCysGlnLeuAspAsnAla 160
DB 421 TCCAGTCCCGAATCCGTTATTTAATCTCTTAAACCTGGGGTGGCCAACTGATTAACGGC 480
QY 161 AsnValThrAspAlaIleLeuGlnArgAlaGlyIleGlySerIleAlaAspPheThrGlyHis 180
DB 481 AACGTACTGACGCCGATCTCCAGAGGCGAGAGGGTCAATGTGCTGACTTACCGGGCAT 540
QY 181 TygIlnThrAlaPheArgIleuLeuGlnArgValaLeuAsnPheProGlnSerAsnLeuCys 200
DB 541 TATCAACGCCGCTTTCGCGAATCGGAACGGGCTTAAATTTCCGCAATCAAACTTGTC 600
QY 201 LeuLysArgGlnLysGlnAspLysCysSerLeuThrGlnAlaLeuProSerGlnLeu 220
DB 601 CTTAAACGTGAAGAAACAGACGAAAGCTGTTCAATTAACCGACATTAACATCGGAATCC 660
QY 221 LysValSerAlaAspCysValSerLeuThrGlyValaIleSerLeuAlaSerMetLeuThr 240
DB 661 AAGGTAGCGCCGACTGTGTCTCATTAACCGGTGGGTAAAGCTCGCATCAATGCTGACG 720
QY 241 GluIlePheLeuLeuGlnGlnAlaGlnIleMetProGluProGlyTyrosIleThr 260
DB 721 GAGATATTTCTCTCCACAAACACAGGAAATGCCGAGCCGGGTGGGGAAGCATCAC 780
QY 261 AspSerHisGlyTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuGln 280
DB 781 GATTCAACACCAAGTGAACACTTGTCTAAGTTGATTAACGCCCAATTTGATTTGCACAA 840
QY 281 ArgThrProGlnValAlaAspSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
DB 841 CCGACGCCAGAGATTGCCCGACGCCGCGCACCCGTTATTGATTGATCAAGACAGCG 900
QY 301 LeuThrProHisAspProGlnLysGlnAlaTyrosIleValThrLeuProThrSerValLeu 320
DB 901 TTGACGCCCCCAATCCACCGGAAABACAGGGTATGTGTGATTAACCTTCAATGCTG 960
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyIleValaLeuGlnLeuAsnTrp 340
DB 961 TTTATTCGCGGACAGCATTAATCTGCAAAATTCGGGGGCGCACTGGAGCTCAACTGG 1020
QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlnLeuValPheGlnArgTrp 360
DB 1021 ACGCTTCCCGGTACCCGATTAACAGCCGCGAGGTGTAATCTGTTAAACGCTGG 1080
QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
DB 1081 CGTGGCTTAAGGATTAACAGCAAGTGAATTCAGTTCGCTGCTTCCAGACTTTACAG 1140
QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValaIleLeuThr 400
DB 1141 CAGATGCGGTAAACCCCTGTCATTAATAACGCCGCGAGAGGTGAACTGAC 1200
QY 401 LeuIleGlyCysGlnGluArgAsnAlaGlnIleMetCysSerLeuAlaGlyPheThrGln 420
DB 1201 CTGGCAGATGTAAGACGAAATGCGCAGGCAATGTTGTTGGTGAAGTTTAACGAA 1260
QY 421 IleValaGlnAlaArgIleProAlaCysSerLeu 432
DB 1261 ATCGTGAATGAAGCAGCATACCGGCGTCAAGTTTG 1296
RESULT 4
ID ABE75420
ID ABE75420 standard; DNA; 1308 BP.
AC ABE75420;
XX
XX 23-FEB-2006 (first entry)
XX
XX
XX E. coli B modified apba phytase 819pH5 DNA sequence SEQ ID NO: 9.
XX de; coding sequence; mutant; thermostable; phytase;
XX protein stabilization; pharmaceutical; fertilizer; cramp; cell culture;
KW

XX osteoporosis; osteopathic; food; appa.
XX Escherichia coli K12.
OS Synthetic.
XX Key Location/Qualifiers
FH 1..1296
FT /*tag= a
FT /product= "modified phytase appa"
FT CDS
XX US2005281792-A1.
XX 22-DEC-2005.
XX 01-SEP-2004; 2004US-00933115.
XX 13-AUG-1997; 97US-00910798.
XX 01-MAR-1999; 99US-00259214.
XX 13-APR-1999; 99US-00291931.
XX 25-MAY-1999; 99US-00318528.
XX 25-MAY-2000; 2000US-00580515.
XX 24-MAY-2001; 2001US-0086379.
XX (SHOR/) SHORT J M.
XX (KRET/) KRETZ K A.
XX (GRAY/) GRAY K A.
XX (BART/) BARTON N R.
XX (GARR/) GARRETT J B.
XX (ODON/) O'DONOGHUE E.
XX (BAUM/) BAUM W.
XX (ROBE/) ROBERTSON D E.
XX (ZORN/) ZORNER P.
XX Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, Odonoghue E;
XX Baum W, Robertson DE, Zorner P;
XX WPI; 2006-055022/06.
XX P-PSDB; ABE75421.
XX Formulation useful as dietary supplement for treating, preventing or
XX reversing osteoporosis and bone loss, and preventing muscle cramps,
XX PT comprising one or more polypeptides having phytase activity.
XX Example 2; SEQ ID NO 9; 82pp; English.
XX The present sequence is that of a polynucleotide encoding a novel
XX modified phytase derived from the Escherichia coli appa protein, with
XX improved thermal tolerance and protease stability compared to the wild-
XX type. The present invention relates to a novel formulation, comprising at
XX least one polypeptide having phytase activity, where the polypeptide is a
XX modified sequence (ABE75421) derived from the Escherichia coli K-12 appa
XX phytase (ABE75419). The modification of the enzyme, comprising a series
XX of mutations, improves the thermal tolerance and protease stability of
XX the protein. The specification also claims a pharmaceutical composition;
XX a kit; an immobilized phytase; a fertilizer or soil additive; a liquid
XX supplement for preventing muscle cramps; a hydrating agent; a tissue
XX culture or cell culture media; and a plant food additive, all comprising
XX the novel phytase, and a method of reducing pollution and increasing
XX nutrient availability in an environment or environmental sample by
XX degrading environmental phytic acid. The novel enzyme is a dietary
XX supplement useful for treating, preventing or reversing osteoporosis or
XX bone loss, and preventing muscle cramps. The liquid supplement is useful
XX for preventing muscle cramps. The method is useful for reducing pollution
XX and increasing nutrient availability in an environment or environmental
XX sample by degrading environmental phytic acid, where the environment or
XX environmental sample comprises a soil or a body of water. The immobilized
XX phytase is useful in foodstuffs for improving the feeding value of
XX phytate rich ingredients.
XX Sequence 1308 BP; 315 A; 345 C; 357 G; 291 T; 0 U; 0 Other;
Alignment Scores: 3.47e-202 Length: 1308
Pred. No.:

Score: 2258.00 Matches: 432
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 15 Gaps: 0
US-10-601-319-10 (1-432) x ABE75420 (1-1308)
QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAAGCGATCTTAATCCATCTTTTATCTCTTGATTCGTTAACCCCGCAATCTGCA 60
QY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValValIleValSerArgHisGly 40
Db 61 TTGGCTCAGAGTAGCCGGAGCTGAACCTGGAAGTGGTGAATTCGACGCGCATGGT 120
QY 41 ValArgAlaProThrLysAlaIleThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 121 GTGCGTCTCCAAACCAAGCCAGCCAGCACTGATGAGGATGTCAACCCAGACGCAATGACCA 180
QY 61 ThrTrpProValLysLeuGlnLeuThrProArgLysGlnLeuIleAlaTyrLeu 80
Db 181 ACTGSCCGGTAAACTGGGTGAGCTGACACCGCGGTGGTGAAGCTTAATCGCTTATCTC 240
QY 81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
Db 241 GGAATTACTGGCGGTGACGGCTGTGGTACCGGATGGCTGCTTAATGTGGCTGCCG 300
QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGlnArgThrArgLysGlyGlu 120
Db 301 CAGTCTGTGTCAGGTGCGGATTTGCTGATGTCAACAGCGTACCCTTAACACGGCGAA 360
QY 121 AlaPheAlaIleGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCTTCGCGCGCGGCTGGACCTGACTGTGCAATAACCTGACTACCGACGAGATACG 420
QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCCAGTCCCGATCCGTTATTTAATCTCTAATAAATCGCGTTGCCAATCGATTAACGCG 480
QY 161 AsnValThrAspAlaIleLeuGlnArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGACTGACGGGATCTCGAGAGGGCAGAGGCGTCAATTCGTGCTTTACCGGGCAT 540
QY 181 TyrGlnThrAlaPheArgGlnLeuGlnArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 541 TATTAACGGCGGTTTCGGAACTGGAACGGGTGCTTAATTTCCGCAATCAACTGTGTC 600
QY 201 LeuLysArgGlnLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlnLeu 220
Db 601 CTTAAAGTGAAGAAACAGAACGAAAGCTGTTAATCGAGGCAATTAACATCGGAATCC 660
QY 221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGAGCGCGCATGTGCTCTTAACCGGTGCGGTAAACCTCGCATCATCTCAACG 720
QY 241 GlnIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
Db 721 GAGATATTTCTCCGCAACAAAGCAAGGAATGCCGAGCCGGGTGGGAAGATCAAC 780
QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuGln 280
Db 781 GATTCAACCAAGTGAACACTCTTAAGTTGATTAACGGGCAATTTGATTTCTCAAA 840
QY 281 ArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
Db 841 CGCAGCGCAGAGGTTGCCCGGACGGCGCACCCCGTTATTAGATTGATTAACAACAGCG 900
QY 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 901 TTGACGCCCATTCACCGCAAAACAGGCGTATGTGTGATTAACCATTCCTGCTGCTG 960
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTrp 340

```

Db      961 TTTATCGCGGACACGATCTAATCTGCAAACTCGCGCGGCACTGAGACTCAACTG 1020
Qy      341 ThrLeuProGlyGlnProAspAnthrProProGlyGlyGlnLeuValPheGlnuArgTP 360
Db      1021 ACGCTTCCCGGTGACGCCGATTAACAGCGCCGAGGTGAACTGTGTTGAACGCTGG 1080
Qy      361 ArgArgLeuSerAspAspSerGlnTPPIeglnValSerLeuValPheGlnThrLeuGln 380
Db      1081 CPTCGGCTAAGGATTAACAGCCAGTGATTCAGGTTCCGTGCTTCCAGACTTTACAG 1140
Qy      381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
Db      1141 CAGATCCGCGATTAACAGCCCGCTGTCAATTAAATCCGCCCGAGAGAGTGAAACTGACC 1200
Qy      401 LeuAlaGlyCysGlnGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db      1201 CTGGCGAGATGTGAAGAGCGAAATGCGCAGGCGCATGTGTTGTCGACAGTTTACGCA 1260
Qy      421 lIleValAsnGluAlaArglIleProAlaCysSerLeu 432
Db      1261 ATCGTAATGAAGCAGCATACCGCGGTGCAAGTTTG 1296

RESULT 5
AEE75422
ID      ABE75422 standard; DNA; 1901 BP.
XX
XX      ABE75422;
AC
XX      23-FEB-2006 (first entry)
DT
XX
XX      E. coli B modified appa phytase 819PH59 DNA sequence SRQ ID NO: 9.
DE
XX
XX      de; gene; mutant; thermostable; phytase; protein stabilization;
KW      pharmaceutical; fertilizer; cramp; cell culture; osteoporosis;
XX      osteopathic; food; appa.
XX
OS      Escherichia coli K12.
OS
XX      Synthetic.
FH
XX      Key
FH      CDS
FT      Location/Qualifiers
FT      188..1486
FT      /tag= a
FT      /product= "modified Appa phytase"
FT      /replace= (389,T)
FT      /tag= b
FT      /note= "Wild type base T replaced by G"
FT      /replace= (390,G)
FT      /tag= C
FT      /note= "Wild type base G replaced by A"
FT      401..403
FT      /tag= d
FT      /note= "ecodes Arg"
FT      /replace= (437,C)
FT      /tag= e
FT      /note= "Wild type base C replaced by T"
FT      /replace= (438,A)
FT      /tag= f
FT      /note= "Wild type base A replaced by G"
FT      /replace= (439,A)
FT      /tag= g
FT      /note= "Wild type base A replaced by G"
FT      /replace= (470,G)
FT      /tag= h
FT      /note= "Wild type base G replaced by C"
FT      /replace= (472,G)
FT      /tag= i
FT      /note= "Wild type base G replaced by T"
FT      /replace= (476,A)
FT      /tag= j
FT      /note= "Wild type base A replaced by T"
FT      /replace= (477,A)
FT      /tag= k

```

```

FT      /note= "Wild type base A replaced by G"
FT      /replace= (478,G)
FT      /tag= l
FT      /note= "Wild type base G replaced by T"
FT      /replace= (689,A)
FT      /tag= m
FT      /note= "Wild type base A replaced by G"
FT      /replace= (690,G)
FT      /tag= n
FT      /note= "Wild type base G replaced by A"
FT      /replace= (691,C)
FT      /tag= o
FT      /note= "Wild type base C replaced by G"
FT      /replace= (728,C)
FT      /tag= p
FT      /note= "Wild type base C replaced by T"
FT      /replace= (729,G)
FT      /tag= q
FT      /note= "Wild type base G replaced by A"
FT      /replace= (730,G)
FT      /tag= r
FT      /note= "Wild type base G replaced by T"
FT      /replace= (863,A)
FT      /tag= s
FT      /note= "Wild type base A replaced by T"
FT      /replace= (864,A)
FT      /tag= t
FT      /note= "Wild type base A replaced by G"
FT      /replace= (1016,T)
FT      /tag= u
FT      /note= "Wild type base T replaced by G"

US2005281792-A1.
XX
XX      22-DEC-2005.
XX
XX      01-SEP-2004; 2004US-00933115.
XX
XX      13-AUG-1997; 97US-00910798.
XX      01-MAR-1999; 99US-00259214.
XX      13-APR-1999; 99US-00291931.
XX      25-MAY-1999; 99US-00318528.
XX      25-MAY-2000; 2000US-00580515.
XX      24-MAY-2001; 2001US-00866379.
XX
XX      (SHOR/) SHORT J M.
XX      (KRET/) KRETZ K A.
XX      (GRAY/) GRAY K A.
XX      (BART/) BARTON N K.
XX      (GARR/) GARRETT J B.
XX      (ODON/) ODOMOGHUE E.
XX      (BAUM/) BAUM W.
XX      (ROBE/) ROBERTSON D E.
XX      (ZORN/) ZORNER P.
XX
XX      Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, Odomoghue E;
XX      Baum W, Robertson DE, Zorner P;
XX      WPI: 2006-055022/06.
XX      P-PSDB; ABE75421.
XX
XX      Formulation useful as dietary supplement for treating, preventing or
XX      reversing osteoporosis and bone loss, and preventing muscle cramps,
XX      comprises one or more polypeptides having phytase activity.
XX
XX      Claim 1; Page; 82pp; English.
XX
XX      The present sequence is that of a polynucleotide encoding a novel
XX      modified phytase derived from the Escherichia coli appa protein, with
XX      improved thermal tolerance and protease stability compared to the wild-
XX      type. The present invention relates to a novel formulation, comprising at
XX      least one polypeptide having phytase activity, where the polypeptide is a
XX      modified sequence (ABE75421) derived from the Escherichia coli K-12 appa

```

CC phyase (ABE75419). The modification of the enzyme, comprising a series
CC of mutations, improves the thermal tolerance and protease stability of
CC the protein. The specification also claims a pharmaceutical composition;
CC a kit; an immobilized phyase; a fertilizer or soil additive; a liquid
CC supplement for preventing muscle cramps; a hydrating agent; a tissue
CC culture or cell culture media; and a plant food additive, all comprising
CC the novel phyase, and a method of reducing pollution and increasing
CC nutrient availability in an environment or environmental sample by
CC degrading environmental phytic acid. The novel enzyme is a dietary
CC supplement useful for treating, preventing or reversing osteoporosis or
CC bone loss, and preventing muscle cramps. The liquid supplement is useful
CC for preventing muscle cramps. The method is useful for reducing pollution
CC and increasing nutrient availability in an environment or environmental
CC sample by degrading environmental phytic acid, where the environment or
CC environmental sample comprises a soil or a body of water. The immobilized
CC phyase is useful in foodstuffs for improving the feeding value of
CC phytate rich ingredients. NOTE: This sequence is not shown in the
CC specification but was created from the wild-type sequence (ABE75418) and
CC the information given in claim 1.

XX
SQ Sequence 1901 BP; 470 A; 497 C; 500 G; 433 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 5.85e-202 Length: 1901
Score: 2258.00 Matches: 432
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 15 Gaps: 0

US-10-601-319-10 (1-432) x ABE75422 (1-1901)

Qy 1 MetLysAlaIleLeuIIleProPheLeuSerLeuLeuIIleProLeuThrProGlnSerAla 20
Db 188 ATGAAGGAGCTTAAATCCCATTTTATCTTTCGATTCGGTAAACCCCGCAATCTGCA 247
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40
Db 248 TTCGCTCAGAGTGAAGCCGAGAGCTGAAGTGGTGAATGGTATGTCAGTCGTCATGT 307
Qy 41 ValArgAlaProThrLysAlaIleThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 308 GTGGGTGCTCCAAACCAAGGCCAGCACTGATGTCAGATGTCACCCAGACGCAATGGCCA 367
Qy 61 ThrTrpProValLysLeuGlyGlnLeuThrProArgGlyGlyGlnLeuIleAlaTyrLeu 80
Db 368 ACCTGGCCGGTAAACTGGGTGAAGCTGACCCGGCGGTGGTGAAGCTAATGCGCTATCTC 427
Qy 81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
Db 428 GGAAGTATTCGGCGCAGCGCTGTGTACCGAGATTCGCTTAATGTGGCTGGCCG 487
Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGln 120
Db 488 CAGTCTGGTCAAGTCGCGATTATGCTGATGTCAGCAGCGTAAACCCGTAACACAGGCGAA 547
Qy 121 AlaPheAlaIleGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCTTCGCGCGCGGCTGGCAGCTGACCTGACCAATTAACCAATTAACCAAGCGATTCG 607
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 608 TCCAGTCCCGATCCGTTATTTTAACTCTTAATAAATGCGCTTTGCAACTGATTAACCGC 667
Qy 161 AsnValThrAspAlaIleLeuGlnArgAlaGlyLysSerIleAlaAspPheThrGlyHis 180
Db 668 AACGAGTACGACGCGATCTCGAGAGGCGCAGAGGATCAATTCGATTCGATTCGCGGAT 727
Qy 181 TyrGlnThrAlaPheArgGlnLeuGlnArgValLeuAsnPheProGlnSerHisLeuCys 200
Db 728 TATCAAAAGCGGCTTCGGAACCTGGAACGGGTGCTTAATTTTCCGCAATCAAACTGTGTC 787
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220

Db 768 CTTAAAGTGAAGAAAGAGCAAGAAAGCTGTTCAATTAAAGCAGGCAATTAACATCGAACTC 847
Qy 221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 848 AAGGTGAGCGCGCATGTGTCTCAATTAAACCGGTGCGGTAAAGCCCTGCATCAATCTACG 907
Qy 241 GluIlePheLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Db 908 GAGTATTTTCTCTGCAACAGCACAGGAATGCCAGACCGGGGTGGGAAGATACAC 967
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
Db 968 GATTCAACACCAAGTGAACACCTTGCTAAGTTTGATTAACGGCAATTTGATTCTCAAA 1027
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
Db 1028 CGCAGCGCAGAGGTTGGCCGCGAGCGCGCACCCGTTATTAGATTGAATCAACAGACGCG 1087
Qy 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 1088 TTGAAGCCCAATCCACCGCAAAACAGGCGTATGTGTGAACATTAAACCACTTCAGTCTG 1147
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGlnLeuAsnTrp 340
Db 1148 TTTATGCGCGGACACGATACCTAATCTGGCAATCTGGCGCGGACCTGAGCTCAACTGG 1207
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlnLeuValPheGluArgTrp 360
Db 1208 ACGCTTCCCGTCCGCGCGGATTAACACCGCCGAGGTGTGAACGTGGTGAACGCTGG 1267
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1268 GGTGGCTAAGCATTAAACAGCCAGTGAATTCAGGTTTCGCTGCTTCCAGACTTTACAG 1327
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlnValLysLeuThr 400
Db 1328 CAGATGGGTGATTAAGCCGCGCTGTCTAATTAATCCGCGCGGAGAGGTAAGTAACCTAC 1387
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1388 CTGGCAGAGATGTGAAGGCAAAATGCGCAGGCGCATGTGTCTGTGGCAGGTTTACGCAA 1447
Qy 421 IleValAsnGlnAlaArgGlyLeuProAlaCysSerLeu 432
Db 1448 ATCGTAATGAAGCACCATTAACCGCGGTGCAATTTG 1483
RESULT 6
AD050296 standard; DNA; 1901 BP.
ID AD050296;
AC AD050296;
XX 29-JUL-2004 (first entry)
XX
XX Escherichia coli K12 appa phyase mutant DNA #2.
DE
XX
XX Appa phyase; bacteria; thermal tolerance; protease stability; foodstuff;
KW phyate; animal feed; fish feed; dough; baking; gene; mutant; ds.
XX
XX Escherichia coli; K12.
OS
XX Synthetic.
XX
FH Key Location/Qualifiers
FT mutation replace(437,C)
FT mutation /*tag= a
FT mutation replace(438,A)
FT mutation /*tag= b
FT mutation replace(439,A)
FT /*tag= c
XX
XX US2004091968-A1.

PD 13-MAY-2004.
 XX 20-JUN-2003; 2003US-00601319.
 PF 13-AUG-1997; 97US-00910798.
 XX 01-MAR-1999; 99US-00259214.
 PR 13-APR-1999; 99US-00291931.
 PR 25-MAY-1999; 99US-00318528.
 PR 25-MAY-2000; 2000US-00580515.
 PR 24-MAY-2001; 2001US-00866379.
 XX (SHOR/) SHORT J M.
 PA (KRETT/) KRETTZ K.
 PA (GRAY/) GRAY K A.
 PA (BART/) BARTON N R.
 PA (GARR/) GARRETT J B.
 PA (ODON/) O'DONOGHUE E.
 PA (MATH/) MATHER E J.
 PI Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E,
 PI Mather EJ;
 XX MPI; 2004-374952/35.
 DR Producing phytase, involves providing nucleic acid derived from bacteria
 PT encoding polypeptide having phytase activity, and expressing nucleic acid
 PT in yeast.
 XX Disclosure; SEQ ID NO 6; 74bp; English.
 XX The invention relates to a method for producing phytase that involves
 CC providing a nucleic acid encoding phytase derived from a bacteria, and
 CC expressing the nucleic acid in a yeast under conditions that allow
 CC expression of the enzyme in the yeast. The invention also relates to
 CC modified phytase enzyme which has improved thermal tolerance and protease
 CC stability at low pH. The phytase enzyme can be used in feedstuffs to
 CC improve the feeding value of phytate rich ingredients, and in diet of
 CC numerous animals including mammals, fowls and fishes, commercially
 CC significant mammals such as pigs, goats, laboratory rodents, commercially
 CC commercially farmed fish such as chicken, ducks, doves, parrot, etc.,
 CC etc., in dough making and baking, in dietary aids for animals, the method
 CC provides easy manufacture of the active ingredient loaded biocompatible
 CC composition, higher yields and loading efficiency. The phytase
 CC incorporated in to the dietary aid is safe for animals. The present
 CC sequence is Escherichia coli K12 apba phytase mutant DNA.
 XX
 SQ Sequence 1901 BP; 472 A; 498 C; 501 G; 429 T; 0 U; 1 Other;
 Alignment Scores:
 Pred. No.: 4,9e-196 Length: 1901
 Score: 2195.00 Matches: 425
 Percent Similarity: 98.44 Conservative: 0
 Best Local Similarity: 98.44 Mismatches: 7
 Query Match: 97.24 Indels: 0
 DB: 12 Gaps: 0
 US-10-601-319-10 (1-432) x ADO50296 (1-1901)
 QY 1 Methylalalleuileuprophenusserleuileuproleuthrprogluinsera1a 20
 DB 188 ATGAAGAGGATCTTAATCCATTTTATCTCTTCGATTCGTTAACCCCGCAATCTTCA 247
 QY 21 Phealaginsergluprogluileuylseuileuinservalvallevalserarghisgl 40
 DB 248 TTCGCTTCAGAGTGAGCCCGAGAGCTGAAAGTGGAGTGTGATGTGATGTCATGTCATGCT 307
 QY 41 Valargalaprothrlvalalethrqlinleuwerghnaapvalthrproaspalatrpro 60
 DB 308 GTGGCTGCTTCCAAACCAAGCCACGCACTGATGACGAGTGTCAACCCGAGAGCAGTGGCA 367
 QY 61 Thrtrprovalylseuileuileuileuthrproargllyglileuileualatyrleu 80

DB 368 ACCGCGCCGGTAAACTGGGTTGGCTGACACCGCGNGGTGTGAGCTAATCGCTATCTC 427
 QY 81 Glyhietrrprrargglinargleuvala1aaprglyleuileuprolysegllycyspro 100
 DB 428 GGACATTAATGCGCGCAGCGCTGTGTGAGCGAGATTCCTGCGCAAAAAGGCGTCCCG 487
 QY 101 Glnserglylnvala1alleilealaapvalaapgluargthrarghsthrnglyglu 120
 DB 488 CAGTCTGGTCAGGTGCGCATTTTGTCTGATGTGACAGACGTCATCCGTTAAACAGCGCNA 547
 QY 121 Alahealalaglyleuualaproapcyasalierthvala1sthrnglna1aapthr 140
 DB 548 GCCTTCGCGCGCGGCTGGCACTGATGTGCAATTAACGTCATTAACGAGCAGATACG 607
 QY 141 SerSerProaspProleupheanProleuylsthrnglyvalCysGlnleuaplanala 160
 DB 608 TCCAGTCCCGATCCGTTATTTAATCTCTTAATAAATCGCGGTTGCAACTGATTAACGCG 667
 QY 161 AenValThrAspAla1leuileuileuarglualaglylsertleala1aapPheThrghs 180
 DB 668 AACGTGACTGACGCGATCTTCAGCAGGCGAGAGGTCATATTGACTTAAACCGGCGCAT 727
 QY 181 TyrGlnThrAlaPhearggluleuileuargvalleuAenPheProGlnSerAsnleuCyS 200
 DB 728 CGGCAAAAGGGGTTTCGGGAAGTGGACGGGTCTTAATTTCCGCAATCAACTTGTGC 787
 QY 201 LeuylsargglulysglinaapgluSerCysserleuthrGlnAlaLeuprosergluleu 220
 DB 788 CTTAAACGTGAGAAACAGACGAAAGCTGTTCATTAAACGAGGCACTTCAATCGAATCTC 847
 QY 221 LysValSerAlaAspCysValSerleuthrGlyAlaValSerleuAlaSerMetleuthr 240
 DB 848 AAGGTGACGCGCGCAATATCTCATTAACCGGTGGTGAAGCTGCAATGCTGACG 907
 QY 241 GluilePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
 DB 908 GAGATATTTCTTCCTGCAACAAGCAGAGGAATGCCGAGCCGGGTGGGAAAGATCAC 967
 QY 261 AspSerHisGlnThrPenthrleuileuSerleuHisAsnAlaGlnPheAspLeuLeuGln 280
 DB 968 GATTCAACACCAAGTGAACACCTTCCTAAGTTGCTAATACGCGCAATTTATTTGCTACA 1027
 QY 281 ArgThrProGluValAlaArgSerArgAlaThrProleuLeuAspLeuIleyleThrAla 300
 DB 1028 CGCAGCGCAGAGGTTGCCCGCAGCGCGCCACCCCGTATTAGATTGATCAAGCAGCG 1087
 QY 301 LeuthrProHisProProGlnAlaGlnAlaTyrGlyValThrleuProThrSerValleu 320
 DB 1088 TTGAAGCCCCATCCACCGCAAAAACAGCGCTATGTGTGACATTAACCACTTCAGTGTCTG 1147
 QY 321 Pheilealaglyh1asprThrAsnleuAlaAsnleuGlyglialaleuileuAsnTrp 340
 DB 1148 TTTATCGCGGACACGATTAATCTGGCAATATCTCGCGCGGCACTGGAACCTCAACTCG 1207
 QY 341 ThrleuProGlyGlnProAspAsnThrProProGlygllyglileuvalPheGlnArgTrp 360
 DB 1208 ACGCTTCCCGGTGACCGGATTAACACCGCGCAGGTGTGAACTGTGTGAAACGCTGG 1267
 QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerleuValPheGlnThrleuGln 380
 DB 1268 GTCGCGCTAAGCATTAACGCAAGTGAATTCAGGTTCGCTGTGCTTCACACTTTACAG 1327
 QY 381 GlnMerArgAspLysThrProleuSerleuAsnThrProProGlygllyvalyleuThr 400
 DB 1328 CAGATGCGTGAATAAAGCGCGCTGTCAATTAATACGCGCGCGAGAGGTAAACTGAC 1387
 QY 401 LeuAlaGlyCysgluGluArgAsnAlaGlnGlyMetCysSerleuAlaGlyPheThrGln 420
 DB 1388 CTGGCAGAGATGTGAAGACGAAATGCGCAGGCGCATGTGTCTGTGGCAGGTTTATACGCA 1447
 QY 421 IleValAsnGlnAlaArgIleProAlaCysSerleu 432
 DB 1448 ATCGTGAATGAACGACGATACCGGCGGTGCAAGTTTG 1483

RESULT 7
AD050295
ID AD050295 standard; DNA, 1901 BP.
XX
XX AC AD050295;
XX
XX 29-JUN-2004 (first entry)
XX
XX Escherichia coli K12 appa phytase mutant DNA #1.
XX
XX Appa phytase; bacteria; thermal tolerance; protease stability; foodstuff;
XX phytase; animal feed; fish feed; dough; baking; gene; mutant; ds.
XX
XX Escherichia coli; K12.
OS Synthetic.
OS
XX Key Location/Qualifiers
XX mutation /tag= a
XX mutation replace(389,T)
XX mutation /tag= b
XX
XX US2004091968-A1.
XX
XX 13-MAY-2004.
XX
XX 20-JUN-2003; 2003US-00601319.
XX
XX 13-AUG-1997; 97US-00910798.
XX 01-MAR-1999; 99US-00259214.
XX 13-APR-1999; 99US-00291931.
XX *25-MAY-1999; 99US-00318528.
XX 25-MAY-2000; 2000US-00580515.
XX 24-MAY-2001; 2001US-00866379.
XX
XX (SHOR/) SHORT J M.
XX (KRETT/) KRETT K.
XX (GRAY/) GRAY K A.
XX (BART/) BARTON N R.
XX (GAR/) GARRETT J B.
XX (ODON/) O'DONOGHUE E.
XX (MATH/) MATHER E J.
XX
XX Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
XX Mather EJ;
XX
XX WPI; 2004-374952/35.
XX
XX Producing phytase, involves providing nucleic acid derived from bacteria
XX encoding polypeptide having phytase activity, and expressing nucleic acid
XX in yeast.
XX
XX Disclosure; SEQ ID NO 5; 74pp; English.
XX
XX The invention relates to a method for producing phytase that involves
XX providing a nucleic acid encoding phytase derived from a bacteria, and
XX expressing the nucleic acid in a yeast under conditions that allow
XX expression of the enzyme in the yeast. The invention also relates to
XX modified phytase enzyme which has improved thermal tolerance and protease
XX stability at low pH. The phytase enzyme can be used in foodstuffs to
XX improve the feeding value of phytate rich ingredients, and in diet of
XX numerous animals including mammals, fowls and fishes, commercially
XX significant mammals such as pigs, goats, laboratory rodents, commercially
XX significant avian species such as chicken, ducks, doves, parrot, etc.,
XX commercially farmed fish such as guppy, zebrafish, molly, swordtail,
XX etc., in dough making and baking, in dietary aids for animals. The method
XX provides easy manufacture of the active ingredient loaded biocompatible
XX composition, higher yields and loading efficiency. The phytase
XX incorporated in to the dietary aid is safe for animals. The present
XX sequence is Escherichia coli K12 appa phytase mutant DNA.
XX
XX Sequence 1901 BP; 475 A; 499 C; 499 G; 427 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 1,45e-195 Length: 1901
Score: 2190.00 Matches: 425
Percent Similarity: 98.4% Conservative: 0
Best Local Similarity: 98.4% Mismatches: 7
Query Match: 97.0% Indels: 0
DB: 12 Gaps: 0
US-10-601-319-10 (1-432) x AD050295 (1-1901)
QY 1 MetIysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
DB 188 ATGAAGAGCATCTTAATCCCATTTTATCTCTTGATTCGTTAACCCCGCAATCTGCA 247
QY 21 PheAlaGlnSerGluProGluLeuLeuGlnSerValValIleValSerAlaGly 40
DB 248 TTCCTCGAGTGAAGCCGAGAGCTGAAGCTGAAGAGTGGATGATTCATGTCATAGGT 307
QY 41 ValArgAlaProThrIysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
DB 308 GTGCGTGTCTCAACCAAGGCCACCGCACTGATGAGATGACACCCGAGACGATGGCA 367
QY 61 ThrTrpProValIysLeuGlyGlnLeuThrProArgGlyGlyGlnLeuIleAlaTyrLeu 80
DB 368 ACTGCGCGGTAAACTGGGTGAGCTGCACCGCGGTGGTGGATGATGCTATCTC 427
QY 81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProIysCysGlyCysPro 100
DB 428 GACATTTACCAAGCCGAGCTGTGTGAGCGAGATGTGCGGCAAAAGGCTGCCG 487
QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGlnArgThrArgIysThrGlyGlu 120
DB 488 CAGTCTGTCAAGTTCGCAATTATTCGTGATGTCACAGACGCTACCCGTAACAGGGCA 547
QY 121 AlaPheAlaAlaGlyLeuValaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
DB 548 GCCTTCGCGCGGCGGTGACCTGACCTGATCAATTAACCGTACATCCACGAGATACG 607
QY 141 SerSerProAspProLeuPheAsnProLeuIysThrGlyValCysGlnLeuAspAspAla 160
DB 608 TCCAGTCCCGATCCGTTATTTAATCTCTTAAATCTGCGGTTTCCCAATGAAATCCG 667
QY 161 AsnValThrAspAlaIleLeuGlnArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
DB 668 AACGTGACTGACCGGATCTTCAGCAGGCGACAGAGGCTCAATTCGACTTTACCGGAT 727
QY 181 TyrGlnThrAlaPheArgGlnLeuGlnArgValLeuAsnPheProGlnSerAsnLeuCys 200
DB 728 CGGCAAACGGGCTTCGCAACTGGAACTGGAGCGGCTTAATTTCCGCAATCAAACTTGTGC 787
QY 201 LeuIysArgGlyGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlnLeu 220
DB 788 CTTAAACCTGAGAAACAGAGCAAGAAAGCTTCATTAAACGAGGCAATTACATCGAACTC 847
QY 221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
DB 848 AAGGTGACCGCCGACAAATCTCATTAACCGGTGGGTAAGCTTCGATCATGTGACG 907
QY 241 GluIlePheLeuLeuGlnGlnAlaGlnIlyMetProGluProGlyIlyArgIleThr 260
DB 908 GAGATATTTCTTCGCAACAAGCACAGGGAATGCGGAGCGGGGTGGGAAAGATCACCC 967
QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
DB 968 GATTCACACCGAGTGAAACACTTGTGTAAGTTGCAATACGGGCAATTTATTTGGTACAA 1027
QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleGlyThrAla 300
DB 1028 CGCAGCGCAAGAGTGGCCGCGAGCGCCGACCCGTTATTAAGATTGATCAAGCACAGC 1087
QY 301 LeuThrProHisProGlnIlySerGlnAlaTyrGlyValThrLeuProThrSerValLeu 320

Db 1088 TTGACGCCCATCCACCGCAAAAACAGCGGTATGTGTGACATTACCACTTCAGTGTG 1147
 Qy 321 PheIIeAlaGlyHIshpThrLeuAlaLeuGlyGlyAlaLeuGlyLeuAlaLeuThr 340
 Db 1148 TTATTCGCGGACACATCTATCTGCAAAATCTGGCGGCGCACTGAGCTCAACTG 1207
 Qy 341 ThrLeuProGlyGlyIleProAspAsnThrProProGlyGlyGlyLeuValPheGlyIleThr 360
 Db 1208 AGCTTCCCGGTACAGCCGATACACAGCCGCGAGGTGTGTAAGTGTGTTAAACGCTG 1267
 Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 Db 1268 CCTCGGCTAACGATACAGCCAGTGTGATTCAGTTCCTGCTTCCAGACTTACAG 1327
 Qy 381 GlnMetArgAspLeuThrProLeuSerLeuAlaThrProProGlyGlyValValLeuThr 400
 Db 1328 CAGATGCGGTAAACGCCGCTGCTATTAAATACCCGCCGAGAGGTGAACCTGACC 1387
 Qy 401 LeuAlaGlyCyGlyGlyIleArgAsnAlaGlnGlyMetCySerLeuAlaGlyPheThrGln 420
 Db 1388 CTGGCAGATGTGTAAGACGCAATGCGCAGGCGATGTGTCGTTGCGACAGTTTACGCA 1447
 Qy 421 IleValAsnGlyAlaArgIleProAlaCySerLeu 432
 Db 1448 ATCGTGAATGAAGCAGCATACCGGCGTCAAGTTTG 1483

RESULT 8

ABE75416
 ID ABE75416 standard; DNA; 1901 BP.

AC ABE75416;

DT 23-FEB-2006 (first entry)

DE E. coli phytase related DNA sequence SEQ ID NO: 5.

KW de; thermostable; phytase; protein stabilization; pharmaceutical; fertilizer; cramp; cell culture; osteoporosis; osteopathic; food.

OS Escherichia coli.

PN US2005281792-A1.

PD 22-DEC-2005.

PF 01-SEP-2004; 2004US-00933115.

PR 13-AUG-1997; 97US-00910798.

PR 01-MAR-1999; 99US-00259214.

PR 13-APR-1999; 99US-00291931.

PR 25-MAY-1999; 99US-00318528.

PR 25-MAY-2000; 2000US-00580515.

PR 24-MAY-2001; 2001US-00866379.

PA (SHOR/) SHORT J M.

PA (KRET/) KRETZ K A.

PA (GRAY/) GRAY K A.

PA (BART/) BARTON N R.

PA (GAR/) GARRETT J B.

PA (ODON/) ODOMOGHUE E.

PA (BAUM/) BAUM W.

PA (ROBE/) ROBERTSON D E.

PA (ZORN/) ZORNER P.

PI Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, Odomoghue E;

PI Baum W, Robertson DE, Zorner P;

DR MPI, 2006-055022/06.

PT Formulation useful as dietary supplement for treating, preventing or reversing osteoporosis and bone loss, and preventing muscle cramps, comprises one or more polypeptides having phytase activity.

PS Disclosure; SEQ ID NO 5; 82pp; English.

CC The present sequence is a DNA sequence related to the production of the
 CC novel thermostable / protease resistant *Escherichia coli* phytase
 CC described in the specification. The present invention relates to a novel
 CC formulation, comprising at least one polypeptide having phytase activity,
 CC where the polypeptide is a modified sequence (ABE75422) derived from the
 CC *Escherichia coli* K-12 *apph* phytase (ABE75419). The modification of the
 CC enzyme, comprising a series of mutations, improves the thermal tolerance
 CC and protease stability of the protein. The specification also claims a
 CC pharmaceutical composition; a kit; an immobilized phytase; a fertilizer
 CC or soil additive; a liquid supplement for preventing muscle cramps; a
 CC hydrating agent; a tissue culture or cell culture media; and a plant food
 CC additive, all comprising the novel phytase, and a method of reducing
 CC pollution and increasing nutrient availability in an environment or
 CC environmental sample by degrading environmental phytic acid. The novel
 CC enzyme is a dietary supplement useful for treating, preventing or
 CC reversing osteoporosis or bone loss, and preventing muscle cramps. The
 CC liquid supplement is useful for preventing muscle cramps. The method is
 CC useful for reducing pollution and increasing nutrient availability in an
 CC environment or environmental sample by degrading environmental phytic
 CC acid, where the environment or environmental sample comprises a soil or a
 CC body of water. The immobilized phytase is useful in foodstuffs for
 CC improving the feeding value of phytase rich ingredients. NOTE: The
 CC present sequence is shown in the sequence listing but not discussed
 CC elsewhere in the specification.

SQ Sequence 1901 BP; 475 A; 499 G; 427 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 1,45e-195 Length: 1901
 Score: 2190.00 Matches: 425
 Percent Similarity: 98.44 Conservative: 0
 Best Local Similarity: 98.44 Mismatches: 7
 Query Match: 97.04 Indels: 0
 DB: 15 Gaps: 0

US-10-601-319-10 (1-432) x ABE75416 (1-1901)

Qy 1 MetLyAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
 Db 188 ATGAAGAGGATCTTAATCCATTTTATCTCTTCGATTCCGTTAACCCGAACTGCA 247
 Qy 21 PheAlaGlnSerGluProGlyLeuLeuIleGlnSerValIleValSerArgHisGly 40
 Db 248 TTCGCTCAGAGTGAAGCCGAGACTGAACTGAAAGTGTGATGTCATGTCATGATG 307
 Qy 41 ValArgAlaProThrIleValAlaIleThrGlnMetGlnAspValThrProAspAlaTrpPro 60
 Db 308 GTGCGTGTCCAAACCAAGCCACGCACTGATGCAAGATGTCACCCAGACGCAATGGCCA 367
 Qy 61 ThrTrpProValIleLeuGlyGlyLeuThrProArgIleGlyGlyLeuIleAlaTrpLeu 80
 Db 368 AACTGCGCGGTAAACTGGGTGAGCTGACACCGCGNGGTGTGACTTAATGCCATCTCTC 427
 Qy 81 GlyHisIleTrpArgIleArgLeuValAlaAspGlyLeuLeuProIleCySerGlyCyAspPro 100
 Db 428 GACATTAACCAACGCAAGCTGCTGGTACCGACGATGTGCGCAAAAAGGCTGCCG 487
 Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGlyIleArgIleThrGlyGlyLeu 120
 Db 488 CAGTCTGATGAGTGTGCGATTAATGCTGATGTGACGAGCGATACCCGTTAAACAGCGCA 547
 Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCyAlaIleThrValHisThrGlnAlaAspThr 140
 Db 548 GCTTCGCGCGCGCGGCTGCGACCTGATGTCATTAACCGTACATCCACGAGCATGCG 607
 Qy 141 SerSerProAspProLeuPheAsnProLeuIleThrGlyValCyGlnLeuAspAsnAla 160
 Db 608 TCACGTCCCGATCCGATTAATCTCTAAACCTGCGCTTCCCAACTGATTAACGCG 667
 Qy 161 AsnValThrAspAlaIleLeuGlyArgAlaGlyIleSerIleAlaAspPheThrGlyHis 180

[illegible]

KR99066028-A.
 15-DEC-1999.
 25-MAY-1998; 98KR-00018810.
 25-MAY-1998; 98KR-00018810.
 (WOJ-) WOJIN CO LTD.
 Bae HD, Forceburgh CW, Goloben S, Cheng KJ;
 WPI, 2000-645078/62.
 P-PSDB; AAU77775.
 Novel phytase gene, recombinant phytase and usage thereof.
 Claim 1; Fig 2; 10pp; Korean.
 The invention relates to a novel phytase gene, a recombinant phytase gene
 and their uses. This sequence encodes a phytase associated protein,
 described in the invention
 Sequence 1299 BP; 317 A; 345 C; 355 G; 282 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 4.8e-195 Length: 1299
 Score: 2182.00 Matches: 424
 Percent Similarity: 98.1% Conservative: 0
 Best Local Similarity: 98.1% Mismatches: 8
 Query Match: 96.6% Indels: 0
 DB: Gaps: 0
 US-10-601-319-10 (1-432) x ABK12514 (1-1299)
 QY 1 Metlysaalleleuileprophelenseleuileproleuthrproginserala 20
 Db 1 ATGAAAGGAGATCTTAATCCATTTTATCTCTTGATTCGGTAAACCCGCAATCTGA 60
 QY 21 Phealaginsergluproglinuleubysleuginservalailevalserarghisgly 40
 Db 61 TTCGCTCAGAGAGACCCGAGCTGAAGCTGAAAGTGTTGATTTCTCAGTGTCAATGGT 120
 QY 41 Valargalaprothrllysalaathrglneuwetginaservaltlthrpaspalatp 60
 Db 121 GTCGTGCTCCAAACCAAGGCCACGCAACTGATGACGAGATGTCAACCCACGCGATGGCCA 180
 QY 61 Thrtppprovalylleugllygluleuthrpproargllygluleuilealtlryleu 80
 Db 181 ACCTGCGCGGTAAACCTGGTGGCTGGACACCGCGGGTGTGATGATTAATCGCTATCTC 240
 QY 81 Glyhisrlytrparyrglnargleuvalalaaspglyleuleuprollysacysglycyspro 100
 Db 241 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGAGATTCCTGCGCGAAAAGGCTGCCG 300
 QY 101 Glnserglyglinalvalalailelealeaaservalaaspgluargthrarglythrargly 120
 Db 301 CAGCTCGGGTCAAGTGGCGCATTTATGCTGATGTGTCAGACGACGATCCGTAACACAGCGAA 360
 QY 121 Alaephalaalaglyleualaaproasercysalallethrvalhisthrginalaaspt 140
 Db 361 GCGTTTCGCGCGCGGCTGGCACCTGACTGTGCAATTAACGATATACCCAGCGAGATACG 420
 QY 141 Serseerproasp 160
 Db 421 TCCAGTCCCGATCCGTTATTTATCTCTTAACAACTGGCGTTTCCCACTGGATTAACGCG 480
 QY 161 Asnvalaltnaspaalileleugluargalaglygliserilealaaspphetrglyhis 180
 Db 481 AACGGACGTGAAGCGCATCTCTGACGAGGACGAGGAGGTCAATTCTGACTTACCGGCGAT 540
 QY 181 Tyrglntlrpalapheargylleugluargvalleuasnpheproglinservaleu 200

```

Db      541 CGGCAAGGCGGTTTGGCACTGGAACGGGCTTAATTTCCGCAATCAACTTGCG 600
Qy      201 LeuLYAAGGluLYeGlnaApGluSerYsSerLeuThrGlnaLeuProSerGluLeu 220
Db      601 CTTAAACGTGAAGAACAGACGAAAGCTGTTCACTTAACGAGCATTAACATCGGAATC 660
Qy      221 LysValSerAlaAspCysValSerLeuThrGlnaValSerLeuAlaSerMetLeuThr 240
Db      661 AAGGTGAGCGCGCAATATCTCATTAACCGGTGGTAAAGCTCCGATCAATGCTGACG 720
Qy      241 GluIlePheLeuLeuGlnGlnaGlnGlyMetProGluProGluYTrpGlyArgIleThr 260
Db      721 GAGATATTTCTCTCTGCAACAGACAGGAATCCCGAGCCGGGGTGGGAAGATCAC 780
Qy      261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuGln 280
Db      781 GATTCAACACAGTGGAAACACTTGTGTAAGTTGCAAAACGCCAATTTATTATGCAACA 840
Qy      281 ArgThrProGluValAlaAspSerArgAlaThrProLeuLeuAspLeuIleYsThrAla 300
Db      841 CGCAGCGCCAGAGTTGCCCGCAGCGCGCCACCCCTGTATGTAGATTGTCAAGACAGCG 900
Qy      301 LeuThrProHisProProGluInLYeGlnaIaYrGlyValThrLeuProThrSerValLeu 320
Db      901 TTGACCGCCCATCCACCGCAAAACAGCGGTATGTGTGACATTACCCACTTCAGTCTG 960
Qy      321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyValaLeuGluLeuAsnTrp 340
Db      961 TTTATCGCCGACACAGTACTATATCTGGCAATCTCGGGCGGCACTGGAAGCTCAACTG 1020
Qy      341 ThrLeuProGlyValProAspAsnThrProProGlyValGluLeuValPheGluArgTrp 360
Db      1021 ACGCTTCCCGGTCACAGCGGATACACGCCCGCAGGGTGAATCTGTGTTGAACGCTTG 1080
Qy      361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db      1081 CCGTCGGCTAAGCGATACAGCGAGTGGATTCCAGTTCCCTGCTTCCAGACTTTACG 1140
Qy      381 GlnMetArgAspLYeThrProLeuSerLeuAsnThrProProGlyGluValLYeLeuThr 400
Db      1141 CAGATCGCGTGAATAACCGCGCTGTCAATTAATACGCCCGCGGAGAGTGAACCTGACC 1200
Qy      401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db      1201 CTGGCAGGATGTGAAGAGGAATGCGCAGGCGCATGTGTTGTCGACAGTTTTCGCA 1260
Qy      421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db      1261 ATCGTGAATGAAGCAGCATACCGCGTGCAGTTTG 1296

```

```

PR      25-OCT-2001; 2001US-0342923P.
PR      08-FEB-2002; 2002US-00072851.
PR      06-MAR-2002; 2002US-0362699P.
XX
XX      (ELITR-) ELITRA PHARM INC.
PI      Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW;
PI      Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX      WPI; 2003-029926/02.
DR      P-PSDB; ABU15427.
XX
XX      New antisense nucleic acids, useful for identifying proteins or screening
PT      for homologous nucleic acids required for cellular proliferation to
PT      isolate candidate molecules for rational drug discovery programs.
XX
XX      Claim 14; SEQ ID NO 7167; 1766pp; English.
XX
CC      The invention relates to an isolated nucleic acid comprising any one of
CC      the 6213 antisense sequences given in the specification where expression
CC      of the nucleic acid inhibits proliferation of a cell. Also included are:
CC      (1) a vector comprising a promoter operably linked to the nucleic acid
CC      encoding a polypeptide whose expression is inhibited by the antisense
CC      nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC      polypeptide or its fragment whose expression is inhibited by the
CC      antisense nucleic acid; (4) an antibody capable of specifically binding
CC      the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC      proliferation or the activity of a gene in an operon required for
CC      proliferation; (7) identifying a compound that influences the activity of
CC      the gene product or that has an activity against a biological pathway
CC      required for proliferation, or that inhibits cellular proliferation; (8)
CC      identifying a gene required for cellular proliferation or the biological
CC      pathway in which a proliferation-regulated gene or its gene product lies
CC      or a gene on which the test compound that inhibits proliferation of an
CC      organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC      compound's activity; (11) a culture comprising strains in which the gene
CC      product is overexpressed or underexpressed; (12) determining the extent
CC      to which each of the strains is present in a culture or collection of
CC      strains; or (13) identifying the target of a compound that inhibits the
CC      proliferation of an organism. The antisense nucleic acids are useful for
CC      identifying proteins or screening for homologous nucleic acids required
CC      for cellular proliferation to isolate candidate molecules for rational
CC      drug discovery programs, or for screening homologous nucleic acids
CC      required for proliferation in cells other than S. aureus, S. typhimurium,
CC      K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC      CC prokaryotic essential genes. Note: The sequence data for this patent did
CC      not form part of the printed specification, but was obtained in
CC      electronic format directly from WIPO at
CC      ftp.wipo.int/pub/published_pat_sequences
XX
XX      SQ Sequence 1299 BP; 317 A; 345 C; 355 G; 282 T; 0 U; 0 Other;
XX
XX      Alignment Scores:
XX      Pred. No.: 488-195 Length: 1299
XX      Score: 2182.00 Matches: 424
XX      Percent Similarity: 98.1% Conservative: 0
XX      Best Local Similarity: 98.1% Mismatches: 8
XX      Query Match: 96.6% Indels: 0
XX      DB: 8 Gaps: 0
XX
XX      US-10-601-319-10 (1-432) x ACA19297 (1-1299)
Qy      1 MetLYaAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db      1 ATGAAGGAGACTTATATCCATTTTATCTCTTGATTCGTTACCGTAAACCCGCAATCTGCA 60
Qy      21 PheAlaGlnSerGluProGluLeuLeuLYeLeuGluSerValIleValSerArgHisGly 40
Db      61 TTCGCTCAGAGTGAAGCCGAGACTGAAGCTGAAGATGTGTGATGTCAAGTCGTCATGAT 120
Qy      41 ValArgAlaProThrThyAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db      121 GTGCTGCTTCAACCAAGGCCACGCACTGATGCAAGATGTCACCCCGACGATGGCCA 180

```

QY 61 ThrTrpProValIysLeuGlyGluLeuThrProArgGlyGlyIleuIleAlaTyrIleu 80
 DB 181 ACCGCGCGGTAACCTGGGTTGGCTGACACCGCGGTTGGTGAAGCTATATCTTC 240
 QY 81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
 DB 241 GGACATTACCAACCGCCAGCTCTGGTAGCCGAGATTGCTGGCGAAAAAGGCTGCCG 300
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGlnArgGlnArgGlyThrGlyGlu 120
 DB 301 CAGTCTGTCAGGTCGATTATGCTGATGTCAGCGAGCGTACCCGTAACACGGCGAA 360
 QY 121 AlaPheAlaIleAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 DB 361 GCCCTCCGCGCGGCTGGACCTGACCTGTCATTAACCTGACATACCCAGGAGATACG 420
 QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 DB 421 TCCAGTCCCGATCCGTTATTTAACTCTTAAAACTGGCGTTGGCACTGGATACGCG 480
 QY 161 AsnValThrAspAlaIleLeuGlnArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 DB 481 AACGTGACTGACGGGATCTCAGCGAGGCGAGGAGGTCATGCTGACTTAACTGGGCT 540
 QY 181 TyrGlnThrAlaPheArgGlyLeuGlyLysValLeuAsnPheProGlnSerAsnLeuCys 200
 DB 541 CGGGAACGGCGTTTCGGGAACCTGGAACTGGATCTTAATTTCCGCAATCAAACTTGGC 600
 QY 201 LeuLysArgGlyLysGlnAspGlySerCysSerLeuThrGlnAlaLeuProSerGlyLeu 220
 DB 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCACTTAACGAGGACTTACATCGGAATCTC 660
 QY 221 LysValSerAlaAspCysValSerLeuThrGlyValAlaValSerLeuAlaSerMetLeuThr 240
 DB 661 AAGGTAGCGCGCAATGTCCTCACTTAACCGGTGGGTAACCTGCACTCAATGTCGACG 720
 QY 241 GlnIlePheLeuLeuGlnGlnAlaGlnGlyMetProGlnProGlyLysArgIleThr 260
 DB 721 GAGATATTTCTCTGCAACCAAGCACAGGGAATGCCGAGCGGGGTGGGAAGATCAC 780
 QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
 DB 781 GATTACACACGATGGAACACTTCTTAAGTTGCACTTAACGCGCAATTTATTTCTCTCAA 840
 QY 281 ArgThrProGlyValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
 DB 841 CGCAGCGCAGAGGTTGCCGCGACCGCGCACCCCGTATTAGATTGATCAAGACAGCG 900
 QY 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
 DB 901 TTGACGCCCATTCACCGCAAAAACAGCGATGATGATGATGATGATGATGATGATGATG 960
 QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTrp 340
 DB 961 TTTATCGCCGACACGATACCTAATCTCGCAATCTCGCGCGGCACTGGAAGCTCAACTCG 1020
 QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGlnArgTrp 360
 DB 1021 AGCGTCCCGGTCAGCCGATACACGCGCGAGGTGATGAACTGATTTGAACGCTGG 1080
 QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 DB 1081 CGTGGCGCTAACGATTAACAGCAAGTGATTCAGGTTTCGCTTCCAGACTTTACAG 1140
 QY 381 GlnMetArgAspLysTrpThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
 DB 1141 CAGATGCGCTGATTAACCGCCCTGCTTAATATCGCGCGGAGAGTGAACCTGAC 1200
 QY 401 LeuAlaGlyCysGlnGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 DB 1201 CTGGCAGGATGTGAAGACGAAATCGCAGGCGCATGTGTGCTGGCAGGTTTACGCA 1260

QY 421 IleValAsnGlnAlaArgIleProAlaCysSerLeu 432
 DB 1261 ATCGTGAATGAGCAGCCATACCGCGCTGCAAGTTTG 1296
 RESULT 11
 ID ADL16137 standard; DNA, 1299 BP.
 ID ADL16137;
 AC ADL16137;
 DE 06-MAY-2004 (first entry)
 XX Escherichia coli K-12 AppA phytase gene.
 XX AppA phytase; modified phytase; K-12; animal feed; feed additive;
 KW phosphorus; phytate; myo-inositol hexakisphosphate; food production;
 KW wet grain milling; dental care product; bone resorption; osteoporosis;
 KW osteopathic; kidney stone; metal removal; waste treatment;
 KW cleaning product; rust removal product; gene; ds.
 OS Escherichia coli; K-12.
 FH Key Location/Qualifiers
 FT 1. 1299
 FT CDS /*tag= a
 FT /product= "AppA phytase"
 XX WO2004015084-A2.
 XX 19-FEB-2004.
 PD 11-AUG-2003; 2003WO-US025058.
 PF 12-AUG-2002; 2002US-0403330P.
 PR (GEMV) GENENCOR INT INC.
 XX Blattmann BO, Darzins A, Davis JM, Enzell LP, Morrison TB;
 PI Merchko GT, Schellenberger V;
 DR WPI, 2004-169498/16.
 XX P-PSDB; ADL16129.
 PT Novel recombinant phytase having modified phytase activity comprising
 PT modification of amino acid residues in mature Escherichia coli phytase
 PT designated BB18B2, useful as feed additive.
 XX Claim 19; Fig 21; 107pp; English.
 PS The invention relates to recombinant modified AppA phytases having
 CC modified phytase activity. The AppA phytases are derived from a mature
 CC Escherichia coli AppA phytase designated BB18B2 (see ADL16120), and
 CC comprise substitutions at least one of residues 26, 43, 46, 54, 73, 113,
 CC 126, 184, 228, 384 or 410, or at a residue located within 5 residues of
 CC those listed above. The invention also encompasses polynucleotides
 CC encoding a modified AppA phytase and which further encodes a secretion
 CC signal sequence operable in Bacillus sp.; expression constructs, vectors
 CC and host cells (preferably Bacillus subtilis) comprising the
 CC polynucleotide sequences; and a method for the production of the modified
 CC AppA phytases. The invention also discloses a method of producing an
 CC enzyme having phytase activity; a method of producing a heterologous
 CC polypeptide having phytase activity in Bacillus subtilis; reducing the
 CC level of phosphorus in animal manure; a composition comprising a modified
 CC AppA phytase; and animal feed comprising cells, spores or plant parts,
 CC including seeds, capable of expressing a modified AppA phytase. The
 CC modified AppA phytases are useful as an animal feed additive,
 CC particularly in feeds for non-ruminant livestock such as pigs and poultry
 CC which lack the digestive enzymes for extracting phosphorus from phytate
 CC (myo-inositol hexakisphosphate), a substance found in plant seed which is
 CC an important source of phosphorus. The modified AppA phytases also have
 CC many other agricultural, industrial, medical and nutritional
 CC applications. For example, they can be used in the isolation and
 CC recovery of rare metals to produce lower phosphate homologs of phytate,

CC which may be used in toothpaste and other dental care products as well as
 CC potential treatments or preventatives of bone resorption (e.g., in
 CC osteoporosis) and renal calculi (kidney stones); in food production
 CC (e.g., tofu and nutritional supplements); in wet grain milling; in
 CC cleaning, personal care, medical and nutritional products; in rust
 CC removal products; and in the removal of metals and other pollutants from
 CC such diverse materials as waste products and carbonated beverages. The
 CC present sequence represents *Escherichia coli* strain K-12 *apra* phytase
 CC gene.

XX Sequence 1299 BP; 317 A; 345 C; 355 G; 282 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.8e-195	Length:	1299
Score:	2182.00	Matches:	424
Percent Similarity:	98.1%	Conservative:	0
Best Local Similarity:	98.1%	Mismatches:	8
Query Match:	96.6%	Indels:	0
DB:	12	Gaps:	0

US-10-601-319-10 (1-432) x ADL16137 (1-1299)

QY 1 MetLysAlaIleuLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 DB 1 ATGAAGAAGCATCTTAATCCCATTTTATCTCTGATTCGGTTAACCCGCAATTCGA 60
 QY 21 PheAlaGlnSerGluProGluLeuLysLeuSerValIleValSerArgHisGly 40
 DB 61 TTCGCTCAGAGTGAAGCCGAGCTGGAAGCTGGAAGGTGTGTGTCAGTGTCTAGT 120
 QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
 DB 121 GTGCTGCTCCACCAAGGCCACGCACTGATGCAAGATGTACCCCAACGATGCGCA 180
 QY 61 ThrTrpProValIleLeuLysGlyLeuLeuThrProArgLysGlyLysLeuLeuIleAlaTyrLeu 80
 DB 181 ACCTGGCCGGTAAACTGGGTGGGTGACACCGCCGGTGGTGAAGTATTCCTATCTC 240
 QY 81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
 DB 241 GCACATTCACCAAGCCGCGCTGGTGAAGCCGATTCGTCGCAAAAAGGCTGCGCG 300
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGlnArgThrArgLysThrGlyGln 120
 DB 301 CAGCTCTGATCAGGTGCGCATTAATGCTGATGTCGACGACCGTAACCCGTAACAGCGCA 360
 QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 DB 361 GCCTTCGCGCGCGGCTGSCACTGACTGTGCATTAACCGTACATCCAGCGCATATCG 420
 QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 DB 421 TCCAGTCCCGATCCCGTATTATTCCTTAATAAATCGGGGTTGCAATGATTAACGCG 480
 QY 161 AsnValThrAspAlaIleLeuGlnLysArgLysGlySerIleAlaAspPheThrGlyHis 180
 DB 481 AACGTAAGTGAAGCATCTCTCAGCAGGAGGAGGAGGATTCATTAACGCAATTCGAGCAT 540
 QY 181 TyrGlnThrAlaPheArgGluLeuGlnLysArgValLeuAsnPheProGlnSerAsnLeuCys 200
 DB 541 CGGCAAAACGGGGTTCGGGAACCTGGAACGGGTCTTAATTTCCCAATCAAACTTGTGC 600
 QY 201 LeuLysArgGluLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlnLeu 220
 DB 601 CTTAAACGTGAGAAACAGACGAAAGCTGTTCAATTAACGACGATTAACATCGGAATC 660
 QY 221 LysValSerAlaAspCysValSerLeuThrGlyValIleValSerLeuAlaSerMetLeuThr 240
 DB 661 AAGGTGAAGCGCCGACAAATGTTCTATTAAACGGGTGGGTGAAGCTGCAATCAATGCTGACG 720
 QY 241 GlnIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyLysTrpGlyValArgIleThr 260
 DB 721 GAGATATTCTCTCTCAACAAACAGCACAGGAATCGCGGAGCGGGTGGGGAAGGATCAC 780

QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
 DB 781 GATTACACACAGTGAACACCTTGGTAAGTTGATTAACGCCCAATTTATTTATCTACAA 840
 QY 281 ArgThrProGluValAlaAspSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
 DB 841 CGACGCGCAGAGGTTCGCCGACGCCGCGCACCCGTTATTAATTTGATTAAGACAGCGG 900
 QY 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
 DB 901 TTGAGCGCCCATCCACCGCAAAACAGCGGTATGATGATCACTTAACCACTTCAATGCTG 960
 QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyValAlaLeuGluLeuAsnTrp 340
 DB 961 TTTATTCGCGGACAGATCTTAATCTGGCAAACTCGCGCGGCACTGGAAGCTCAACTGG 1020
 QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyLysLeuValPheGluArgTrp 360
 DB 1021 ACGCTTCCCGGTACCGCATTAACAGCCGCGCAGGTGTGAACCTGTGTTAAGCGCTGG 1080
 QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 DB 1081 CGTCGGCTTAAGGATTAACAGCCAGTGAATTCAGTTTCCTGCTGCTTCCAGACTTACAG 1140
 QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 DB 1141 CAGATGCCGTGATTAACCCCGCTGTCTTAATAATACGCGCCGAGAGGTGAACAGTACC 1200
 QY 401 LeuAlaGlyCysGlnGluAspAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 DB 1201 CTGGCAGAGTGTGAAGAGCGAAATGCGAGGCGCATGTGTCTGTGGCAGGTTTACGCA 1260
 QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
 DB 1261 ATCGTAATGAAGCAGCATTAACCGCGCTGCACTTGG 1296

RESULT 12
 AAD25463 standard; DNA; 1901 BP.
 ID AAD25463;
 AC AAD25463;
 DT 26-MAR-2002 (first entry)
 XX
 DB *Escherichia coli* appa phytase wild type DNA.
 KW Bacterial phytase; K12; appa phytase; protease stability; anabolic;
 KW gastrointestinal; nutritional value; feed treatment process; therapy;
 KW thermal tolerance; growth performance; alcoholic drink; biopulping;
 KW non-alcoholic drink; bioleaching; de.
 OS *Escherichia coli*.
 FH Key Location/Qualifiers
 FT CDS 188..1486
 FT /*tag=a
 FT /product="B. coli appa phytase protein"
 MO200190333-A2.
 PD 29-NOV-2001.
 PF 24-MAY-2001; 2001WO-US017118.
 PR 25-MAY-2000; 2000US-00580515.
 PA (DIVE-) DIVERSA CORP.
 PI Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, O'donoghue E;
 DR WPI; 2002-083108/11.
 DR P-P8DB; AAE15807.

PN US2002136754-A1.
 XX
 PD 26-SEP-2002.
 XX
 PF 24-MAY-2001; 2001US-00866379.
 XX
 PR 13-AUG-1997; 97US-00910798.
 PR 01-MAR-1999; 99US-00259214.
 PR 13-APR-1999; 99US-00291931.
 PR 25-MAY-1999; 99US-00318528.
 PR 25-MAY-2000; 2000US-00580515.
 XX
 XX (SHOR/) SHORT J M.
 PA (KRET/) KRETZ K A.
 PA (GRAY/) GRAY K A.
 PA (BART/) BARTON N R.
 PA (GARR/) GARRETT J B.
 PA (DONO/) O' DONOGHUE E.
 XX (MATH/) MATHUR E J.
 XX
 PI Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB;
 PI O' Donoghue E, Mathur EJ;
 DR P-PSDB; ADA19450.
 DR
 DR WPI; 2003-040002/03.
 XX
 PT Isolated *Escherichia coli* polymnucleotide encoding a modified phytase
 PT enzyme, useful in the production of animal feed, for improving the
 PT nutritional value of phytate-containing feedstuff and for enhancing
 PT digestion in humans and animals.
 XX
 PS Claim 42; Fig 7; 62pp; English.
 XX
 CC The invention relates to an isolated *Escherichia coli* polymnucleotide
 CC encoding a phytase enzyme appearing as ADA19450 and having amino acids
 CC modification from W68B, Q84W, A95P, K97C, S168B, R180Y, N226C or Y277D.
 CC Also included the *E. coli* appa gene ADA19449 (or an oligonucleotide
 CC derived from it) or its mutant sequence ADA19452, expression vectors,
 CC host cells, a method of improving nutritional value of a phytate-
 CC containing feedstuff by contacting the phytate-containing feedstuff with
 CC a pure phytase enzyme mutated phytase (where the phytase enzyme catalyzes
 CC the liberation of inorganic phosphate from the phytate in the phytate-
 CC containing feedstuff), a method to produce an animal feed containing a
 CC microbial phytase (comprising culturing the plant cell, plant part or
 CC plant under conditions where the nucleotide sequence is expressed and
 CC converting the plant cells, plants or plant into a composition for animal
 CC feed), a feed composition for animals (comprising the plant seeds, plant
 CC cells, plant parts or plants in admixture with a phytate-containing
 CC feedstuff), a method to treat a human or an animal able to benefit from
 CC digestive enhancement by the activity of an exogenous phytase enzymes
 CC comprising administering to the human or animal the plant seed, plant
 CC cells, plant parts or plants of a transgenic plant which is modified to
 CC contain an expression system which expresses a nucleotide sequence
 CC encoding a phytase enzyme, a transgenic non-human organism whose genome
 CC comprising a heterologous nucleic acid sequence encoding a polypeptide
 CC having phytase activity. The phytase enzyme is useful for improving the
 CC nutritional value of phytate-containing feedstuff, in the production of
 CC animal feed and for enhancing digestion in humans and animals. The
 CC invented method improves thermal tolerance and protease stability. It
 CC also improves the feeding value of phytate rich ingredients. The present
 CC sequence represents the *E. coli* K12 appa gene encoding wild-type phytase.
 XX
 SQ Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 0 U; 1 Other;
 Alignment Scores:
 Pred. No.: 8, 17e-195 Length: 1901
 Score: 2182.00 Matches: 424
 Percent Similarity: 98.1% Conservative: 0
 Best Local Similarity: 98.1% Mismatches: 8
 Query Match: 96.6% Indels: 0
 DB: 8 Gaps: 0
 US-10-601-319-10 (1-432) x ADA19449 (1-1901)

QY	1	MethylValAlaIleuLeuProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla	20
DB	188	ATGAAAGCATCTTAATCCATTTTATCTCTTGATTCCTGTAACCCCGCAATCTGCA	247
QY	21	PheAlaGlnSerGlnProGlnLeuLeuGlnSerValValIleValSerPheGlnGly	40
DB	248	TTGCTCAAGTGAAGCCGAGCTGAAAGCTGAAAGTGTGATGTCAGTGTCAATGGT	307
QY	41	ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro	60
DB	308	GTGCGGTCTCAACCAAGGCCAAGCACTGATGCAAGAGTCAACCCCAACGATGGCCA	367
QY	61	ThrTrpProValLysLeuGlnLysLeuThrProArgGlyGlyGlnLeuIleAlaTyrLeu	80
DB	368	ACCTGCGCCGTAATCACTGGGTGTGGCTGACACCGCGGTGTGAGCTAATCGCTAATCTTC	427
QY	81	GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro	100
DB	428	GGAATTTACCAAGCCAGCGCTGTGAGCGCAAGATTCTGGCGAATAAGGCTGCGG	487
QY	101	GlnSerGlyGlnValAlaIleIleAlaAspValAspGlnArgThrArgLysThrGlyGln	120
DB	488	CAGTCTGTCAGTCCGATTTATGCTGATGTGACAGACGTAACCCGTAACAGCGCGAA	547
QY	121	AlaPheAlaIleGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr	140
DB	548	GCCTTCGCGCGCGGCTGGCACTGTCGAAATPACCGTACATPACCGACGACGATACG	607
QY	141	SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla	160
DB	608	TCCAGTCCCGATCCGTTATTTAATCTCTTAAACCTGGGTTTCCCACTGATTAACGCG	667
QY	161	AsnValThrAspAlaIleLeuGlnLysArgAlaGlySerIleAlaAspPheThrGlyHis	180
DB	668	AACGTGACGTGACGCAATCTCAGCAGGCGAGAGGCTCAATGCTGACTTACCGGCAAT	727
QY	181	TyrGlnThrAlaPheArgGlnLeuGlnLysValLeuAsnPheProGlnSerAsnLeuCys	200
DB	728	CCGCAAAACCGCTTTCGCGCACTGGAACGGGCTGTTAATTTCCCAATCAAACTTGTC	787
QY	201	LeuLysArgGlnLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlnLeu	220
DB	788	CTTAAACGTGAATAACAGACGAAAGCTGTTATTAAACGACAGCACTTACATCGGAATCC	847
QY	221	LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr	240
DB	848	AAGGTGAGCGCGCAATGTCTCATTAACCGGTGCGTAAGCTCGCATCAATGCTGACG	907
QY	241	GluIlePheLeuLeuGlnGlnAlaGlnLysMetProGlnProGlyTyrGlyValArgIleThr	260
DB	908	GAGATATTTCTCTCTCAACCAAGCAAGGGAATGCGGAGCGGGGTGGGGAAGATCAACC	967
QY	261	AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln	280
DB	968	GATTACACCAATGGAACAACCTTGCTAAGTTGCAATACGCCCAATTTATTGTCTACAA	1027
QY	281	ArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla	300
DB	1028	CCGACGCCAGAGGTGCGCGCAGCGCCGACCCGCTTATTTGATTTGATCAAGACAGG	1087
QY	301	LeuThrProHisAspProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu	320
DB	1088	TTGACGCCCCATTCACCCCAAAAACAGCGGTGTGTGACATPACCACTTCAATGTGCTG	1147
QY	321	PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeuGlnLeuAsnTrp	340
DB	1148	TTTATCGCGGACAGATTAATCTGCAAAATCTGCGGCGGCACTGAGCTCAACTGG	1207
QY	341	ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlnLeuValPheGlnLysTrp	360
DB	1208	ACGCTTCCCGGTGACGCGGATPACACGCGCGAGTGTGAACTGCTTTTGAACGCTGG	1267

```
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1268 CGTGGCGTAAAGCGATTAACAGCCAGTGAATTCAGTTTGGCTGCTTTCAGACTTAAAG 1327
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAnthrProProGlyGluValLysLeuThr 400
Db 1328 CAGATGCGCTGATTAATAACCCGCTGTCATTAAATACGCCGCCGAGAGGTGAACCTGACC 1387
Qy 401 LeuValGlyCysGluGluArgAsnValGlnGlyMetCysSerLeuValGlyPheThrGln 420
Db 1388 CTGGCAGATGTAAGAGCGAAATGCGAGGCGATGTCTGTTGGCAGGTTTACGCA 1447
Qy 421 IleValaengValaArgIleProAlaCysSerLeu 432
Db 1448 ATCGGATGAAGCAGCATACCGCGCGGCGAGTTTG 1483

RESULT 14
ADC87744
ID ADC87744 standard; DNA, 1901 BP.
AC ADC87744;
XX
XX
XX 01-JAN-2004 (first entry)
XX
XX DNA encoding Escherichia coli appa phytase.
XX
XX Phytase; food supplement; enzyme delivery matrix; soybean meal;
XX thermotolerance; thermostability; kernel; phytate;
XX myo-inositol-hexaphosphate; inositol; inorganic phosphate;
XX thermoculant; feed value; digestion; gene; de; appa.
XX
XX Escherichia coli.
XX
XX
XX Key Location/Qualifiers
XX FT 188..1486
XX FT /*tag= a
XX FT /product= "Appa phytase"
XX FT /trans_except= (pos:401..403,aa:Arg)
XX
XX US2003103958-A1.
XX
XX
XX 05-JUN-2003.
XX
XX 24-MAY-2002; 2002US-00156660.
XX
XX 13-AUG-1997; 97US-00910798.
XX 01-MAR-1999; 99US-00259214.
XX 13-APR-1999; 99US-00291931.
XX 25-MAY-1999; 99US-00318528.
XX 25-MAY-2000; 2000US-00580515.
XX 24-MAY-2001; 2001US-0066379.
XX
XX
XX (DIVE-) DIVERSA CORP.
XX
XX Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
XX Mathur EJ;
XX
XX WPI: 2003-787039/74.
XX P-PSDB; ADC87745.
XX
XX New nucleic acid encoding a polypeptide having phytase activity, useful
XX in improving the feeding value of phytate rich ingredients or as an aid
XX in phytate digestion.
XX
XX
XX Example 1, SEQ ID NO 3; 113bp; English.
XX
XX The invention discloses a new isolated or recombinant nucleic acid which
XX encodes a polypeptide having a phytase activity. Also claimed is a
XX nucleic acid probe, an amplification primer sequence pair, an expression
XX cassette comprising the nucleic acid, a vector comprising the nucleic
XX acid, a transgenic non-human animal or plant, or its seed, comprising the
XX nucleic acid, an antisense oligonucleotide, inhibiting the translation of
XX a phytase message in a cell, a heterodimer comprising the polypeptide and
```

```
CC a second domain, an array comprising immobilised polypeptide or nucleic
CC acid, a hybridoma comprising an antibody that specifically binds to the
CC polypeptide, a food supplement for an animal, an edible enzyme delivery
CC matrix, an edible pellet comprising a granule edible carrier and the
CC polypeptide, a feed composition, a soybean meal, isolating or identifying
CC the polypeptide, making an anti-phytase antibody, producing a recombinant
CC polypeptide, determining whether a compound binds to the polypeptide,
CC identifying a modulator, whole cell engineering of new or modified
CC phenotypes by using real-time metabolic flux analysis, increasing
CC thermotolerance or thermostability of the phytase polypeptide, increasing
CC the resistance of the phytase polypeptide to enzymatic inactivation in a
CC digestive system and processing of corn and sorghum kernels. The phytase
CC activity comprises catalysis of phytate (myo-inositol-hexaphosphate) to
CC inositol and inorganic phosphate and is thermotolerant. The nucleic acid
CC is useful in improving the feeding value of phytate rich ingredients or
CC as an aid in phytate digestion. The sequence presented is the DNA
CC encoding the Escherichia coli appa phytase.
CC
XX
XX SQ Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 0 U; 1 Other;
XX
XX Alignment Scores:
XX Pred. No.: 8.17e-195 Length: 1901
XX Score: 2182.00 Matches: 424
XX Percent Similarity: 98.1% Conservative: 0
XX Best Local Similarity: 98.1% Mismatches: 8
XX Query Match: 96.6% Indels: 0
XX DB: 10 Gaps: 0
XX
XX US-10-601-319-10 (1-432) x ADC87744 (1-1901)
Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 188 ATGAAGCGATCTTAATCCATTTTATCTTCTGATTCGTTAACCCGCAATCGCA 247
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValValIleValSerArgHisGly 40
Db 248 TTCCTCAGAGTGAGCCGAGAGCTGAAGCTGGAAGTGTGGATTCAGTCAGTCATGGT 307
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 308 GTGCGTCTCCAAACGAGCCAGCACTGATGCGAGATGTCACCCAGACGATGGCA 367
Qy 61 ThrTrpProValLysLeuGlyGluLeuThrProArgIleGlyGluLeuIleAlaTrpLeu 80
Db 368 ACCTGGCGGTAATACTGGGTGGCTGACACCGGAGTGGTGAATGCTATCTC 427
Qy 81 GlyHisTrpTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
Db 428 GGACATTACCAACCCGAGCGCTGTGTAGCCGACGATTTGCTGGGAAAAAGGCTGCCG 487
Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 488 CAGTCTGTCAGTGGCGATTTATGCTGATGTCAACAGCGCTACCCGTAACACGGCGAA 547
Qy 121 AlaPheAlaIleGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCCCTGGCCCGCGGCGTGGGACCTGACTGTGCAATTAACCGATACACGAGGATACG 607
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 608 TCCAGTCCCGATCCGTTATTTAATCCCTTAATAACGCGCTTTCACCACTGATTAACGCG 667
Qy 161 AsnValThrAspAlaIleLeuGluArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db 668 AACGTGACTAGCGGATCTTCAGCAGGACAGAGGCTCAATGCTGACTTAACGGGCGAT 727
Qy 181 TyrGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 728 CCGGAACGGCGTTTCCGGAAGTGAACGGGTGTTAATTTCCGCAATCAAACTTGTC 787
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 788 CTTAAACGTGAGAAACAGAGCAAGAAAGCTGTTCAATTAAACGACGATTAACATCGGAATC 847
```

```

QY 221 LVVValSerAlaAspCyValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
DB 848 AAGGTAGAGCGCGCAAAATCTCTATTACCGGCGGTAAAGCTCCCAATCAATAGCTGAGG 907
QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGlyProGlyTyrGlyArgIleThr 260
DB 908 GAGATATTCTCTCTCGAACACAGCAGGAAATCCGAGCCGGAGGTGGGAGAGATCACC 967
QY 261 AppSerHisGlnTyrPheThrLeuLeuSerLeuHisAsnLeuGlnPheAspLeuGln 280
DB 968 GATTCAACACAGTGAACACCTTGCTAAGTTTGATTAACCGCAATTTATTTGACAA 1027
QY 281 ArgThrProGlyValAlaIleAspSerArgAlaThrProLeuLeuAspLeuIleYrThrAla 300
DB 1028 CGCAGCGCCAGAGGTGGCCGCGAGCCGCGCACCCCTTTTGTGATTGTGATCAAGACAGG 1087
QY 301 LeuThrProHisProProGlyInLeuGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
DB 1088 TTGAGCGCCCATTCACCGCAAAACAGGCGTATGTGTGATTAACCTTCACTGCTG 1147
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlyLeuAsnTyr 340
DB 1148 TTTATCGCGGACACGATCTATCTGCAAAATCTCGCGGCGCACTGAGGCTCAACTGG 1207
QY 341 ThrLeuProGlyValProAspAsnThrProProGlyValGlyLeuValPheGlyLeuArgTyr 360
DB 1208 AGCTTCCCGGTACCGCGGATTAACGCGCGCAAGGTGGAACTGTGTGTTAAACGCTGG 1267
QY 361 ArgArgLeuSerAspAsnSerGlnTyrIleGlnValSerLeuValPheGlnThrLeuGln 380
DB 1268 COTCGGCTTAAGCGATTAACGAGTGGATTCAGGTTTCCTGTGCTTCCAGACTTTACG 1327
QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
DB 1328 CAGATGCGGATTAACGCGCGCTGTCAATTAACGCGCGCGAGAGGTGAACGTAAC 1387
QY 401 LeuAlaGlyCysGlyLeuArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
DB 1388 CTGGCAGAGATGTGAAGAGGAATGCGCAGGCGCATGTGTGTCGTCAGGTTTACGCA 1447
QY 421 IleValAsnGlyAlaArgIleProAlaCysSerLeu 432
DB 1448 ATCGTAATGAAGCACGCAATACCGGCGTGCAGTTTG 1483
RESULT 15
AD050297
ID AD050297 standard; DNA; 1901 BP.
XX
AC AD050297;
XX
DT 29-JUN-2004 (first entry)
XX
DE Escherichia coli K12 appa phytase DNA.
XX
KM Appa phytase; bacteria; thermal tolerance; protease stability; foodstuff;
XX phytate; animal feed; fish feed; dough; baking; gene; ds.
XX
OS Escherichia coli; K12.
XX
FH Key Location/Qualifiers
FT CDS 188..1486
FT /*tag= a
FT /product= "Appa phytase"
FT unsure 401..403
FT /*tag= b
FT /*note= "Encodes Arg"
XX
PN US2004091968-A1.
XX
PD 13-MAY-2004.
XX
PF 20-JUN-2003; 2003US-00601319.

```

```

XX 13-AUG-1997; 97US-00910798.
PR 01-MAR-1999; 99US-00259214.
PR 13-APR-1999; 99US-00291931.
PR 25-MAY-1999; 99US-00318528.
PR 25-MAY-2000; 2000US-00580515.
PR 24-MAY-2001; 2001US-00866379.
XX
PA (SHOR/) SHORT J M.
PA (KRET/) KRETZ K.
PA (GRAY/) GRAY K A.
PA (BART/) BARTON N R.
PA (GAR/) GARRETT J B.
PA (ODON/) O'DONOGHUE E.
PA (MATH/) MATHER E J.
XX
PI Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E,
PI Mather EJ;
XX
DR WPI; 2004-374952/35.
DR P-PSDB; AD050298.
XX
PT Producing phytase, involves providing nucleic acid derived from bacteria
PT encoding polypeptide having phytase activity, and expressing nucleic acid
PT in yeast.
XX
PS Disclosure; SEQ ID NO 7; 74bp; English.
XX
CC The invention relates to a method for producing phytase that involves
CC providing a nucleic acid encoding phytase derived from a bacteria, and
CC expressing the nucleic acid in a yeast under conditions that allow
CC expression of the enzyme in the yeast. The invention also relates to
CC modified phytase enzyme which has improved thermal tolerance and protease
CC stability at low pH. The phytase enzyme can be used in foodstuffs to
CC improve the feeding value of phytate rich ingredients, and in diet of
CC numerous animals including mammals, fowls and fishes, commercially
CC significant mammals such as pigs, goats, laboratory rodents, commercially
CC significant avian species such as chicken, ducks, doves, parrot, etc.,
CC commercially farmed fish such as guppy, zebrafish, mollly, swordtail,
CC etc., in dough making and baking, in dietary aids for animals. The method
CC provides easy manufacture of the active ingredient loaded biocompatible
CC composition, higher yields and loading efficiency. The phytase
CC incorporated in to the dietary aid is safe for animals. The present
CC sequence is Escherichia coli K12 appa phytase DNA.
XX
SQ Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 0 U; 1 Other;
XX
Alignment Scores:
Pred. No.: 8,17e-195 Length: 1901
Score: 2182.00 Matches: 424
Percent Similarity: 98.14 Conservative: 0
Best Local Similarity: 98.14 Mismatches: 8
Query Match: 96.64 Indels: 0
DB: 12 Gaps: 0
XX
US-10-601-319-10 (1-432) x AD050297 (1-1901)
QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
DB 188 ATGAAGAGCATCTTAATCCATTTTATCTCTTGCATTCGCTTAACCCCGCAATCTGCA 247
QY 21 PheAlaGlnSerGluProGlyLeuLeuLysLeuGlnSerValIleValSerArgHisGly 40
DB 248 TTCGCTCAGAGTGAAGCCGAGCTGAAGCTGAAGAGTGGTGAATGTGATGTCATGTCATG 307
QY 41 ValArgAlaProThrValAlaThrGlnLeuMetGlnAspValIleThrProAspAlaTyrPro 60
DB 308 GTGGGTGCTCCAAACCAAGCCACGCACTGATGAGATGATGACCCCAAGACAGTGGCCA 367
QY 61 ThrTyrProValLysLeuGlyGlyLeuThrProArgGlyGlyGlyLeuIleAlaTyrLeu 80
DB 368 ACCTGCGCGGTAAACTGGTGGCTGACACCGCGNGGTGTGAGCTAATGCGCTATCTC 427

```

Search completed: June 14, 2006, 10:58:28
Job time : 581.385 secs

```
Qy 81 G1YH1STYrTPArGlnArGLeuVal1AaSPGlyLeuLeuProLyCysGlyCysPro 100
Db 428 GGAACATTAACCAACCCGACCGCTCTGGTAGCCACGGAATTGCTGGGAAAAAGGCTGCCG 487
Qy 101 GlnSerGlyGlnVal1Aa1e1Ie1AaSPValASPValArgThrArgLySThrGlyGlu 120
Db 488 CAGTCTGGTCAAGTCGCGCATTAATGCTGATGTCGACGAGCGTACCCGTAACAAACAGCGCAA 547
Qy 121 A1Aahea1Aa1AGlyLeuValAProASPcysA1a1IeThrValH1eThrGlnA1aAspThr 140
Db 548 GCCTTCGCGCGCGCTGGCACTGACTGTCGCAATPACGTAACATCCAGAGCATACG 607
Qy 141 SerSerProASPProLeuPheASPProLeuLySThrGlyValCysGlnLeuASPAsnA1a 160
Db 608 TCCAGTCCCGATCCGTTATTTATCTCTTAATAAATCGGCGTTTGCCAACTGATPACCGCG 667
Qy 161 AsnVal1ThrASPAla1IeLeuGluArgA1AGlyGlySer1IeA1aASPheThrGlyH1s 180
Db 668 AACGTGACTGACGGGATCTCAGCAGGCGACAGAGGGTCAATGCTGACTTAACCGGCGCAT 727
Qy 181 TyrGlnThra1aPheaArgLuleuGluArgValLeuAenPheProGlnSerAsnLeuCys 200
Db 728 CGGCAAAAGCGGCTTCCGAACTGAAAGGAGGCTTAATTTCCGCAATCAAACTTGTC 787
Qy 201 LeuLySAArgGlyLySGlnASPGLuSerCysSerLeuThrGlnA1aLeuProSerGlyLeu 220
Db 788 CTTAAACCTGAAGAAACAGAAAGCTGTTCACTTAACGACGCACTTACATCCGAATC 847
Qy 221 LysVal1SerA1aASPcysVal1SerLeuThrGlyA1aVal1SerLeuA1aSerMetLeuThr 240
Db 848 AAGGTGAGCGCGCAATGTCTCATTTAACCGGTGCGTAAGCTCGCATCAATGCTGACG 907
Qy 241 GlnL1ePheLeuLeuGlnGlnA1AGlnGlyMetProGluProGlyTyrGlyArgGlyLeThr 260
Db 908 GAGATATTCTTCCTGCAACAAAGCAAGGAAATGCCGAGCCGCGGGTGGGAAAGATCACCC 967
Qy 261 ASPSerH1sGlnTPAsnThrLeuLeuSerLeuH1sAenA1AGlnPheASPLeuLeuGln 280
Db 968 GATTCAACACGAGTGAACACTTGTGAAGTTTGCAATAACGCGCAATTTATTGCTACAA 1027
Qy 281 ArgThrProGluVal1AaXSerArgA1aThrProLeuLeuASPLeu1IeLySThrA1a 300
Db 1028 CGCAGCGCCAGAGTGTGCGCGCAGCGCCGCGCATTTATGATCAAGACAGCGC 1087
Qy 301 LeuThrProH1sProProGlnLySGlnA1aTyrGlyVal1ThrLeuProThrSerValLeu 320
Db 1088 TTGACGCCCCCATCCACCGCAAAAACAGCGATAGGTGTGACATTAACCATTCAGTGTCTG 1147
Qy 321 Phe1IeA1AGlyH1sASPThrAsnLeuA1aAsnLeuGlyGlyA1aLeuGluLeuAsnTrp 340
Db 1148 TTTATCGCCGGAACGATACCTAATCTGGCAATCTCGGCGGCGCACTGGAAGCTCAACTCG 1207
Qy 341 ThrLeuProGlyGlnProASPAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 1208 ACGCTTCCCGGTCAAGCCGATPACACGCGCGCAGGTGTGAACGTGTGTGAACGCTCG 1267
Qy 361 ArgArgLeuSerASPAsnSerGlnTrp1IeGlnValSerLeuValPheGlnThrLeuGln 380
Db 1268 CGTGCGCTTAAGCGATPACAGCAAGGATPACAGGTTGTGCTGTTCACAGACTTACAG 1327
Qy 381 GlnMetArgASPlySThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 1328 CAGATGCGGTGATPAAACCGCTGTCACTTAATAACGCGCGCGAGAGGTGAACCTGACC 1387
Qy 401 LeuA1AGlyCysGluGluArgAsnA1AGlnGlyMetCysSerLeuA1AGlyPheThrGln 420
Db 1388 CTGGCAGAGATGTGAAGACGAAATGCGCAGGCGCATGTGTGCTGGCAGTTTTACGCAA 1447
Qy 421 I1IeVal1AsnGluA1aArg1IeProA1aCysSerLeu 432
Db 1448 ATCGAATGAAGACGCAATCCGGCGTGCAGATTG 1483
```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 14, 2006, 14:19:34 ; Search time 1412.42 Seconds

(without alignments)
5637.400 Million cell updates/sec

Title: US-10-601-319-10

Perfect score: 2258
Sequence: 1 KKAIIIPFLSLIPIPLPQSA.....CSLAGPTQIVNARIIPACSL 432

Scoring table:
BLOSUM62
Xgapop 10.0 , Ygapext 0.5
Xgapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame_p2n.model -DEV=xlp
-O=/abs/ABSSMB.spool/US10601319/runat.12062006.180103.27735/app.query.fasta_1
-DB=Published Applications NA Main -QFMT=fastap -SUFFIX=rnpbm -MINMATCH=0.1
-LOOCL=0 -LOOEXT=0 -UNITS=hits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs806p
-USER=US10601319 @CGN 1.1 3524 @runat.12062006.180103.27735 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WALT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA Main:
1: /EMC Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /EMC Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq:*
3: /EMC Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
4: /EMC Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
5: /EMC Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
6: /EMC Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
7: /EMC Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
8: /EMC Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
9: /EMC Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
10: /EMC Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
11: /EMC Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
12: /EMC Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
13: /EMC Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
14: /EMC Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq:*
15: /EMC Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq:*
16: /EMC Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No. Score Match Length DB ID Description

1 2258 100.0 1296 14 US-11-056-354-1 Sequence 1, Appl1

2 2258 100.0 1308 6 US-10-156-660-1 Sequence 1, Appl1
3 2258 100.0 1308 8 US-10-601-319-9 Sequence 9, Appl1
4 2258 100.0 1308 10 US-10-933-115-9 Sequence 9, Appl1
5 2195 97.2 1901 8 US-10-601-319-6 Sequence 6, Appl1
6 2190 97.0 1901 8 US-10-601-319-5 Sequence 5, Appl1
7 2190 97.0 1901 10 US-10-933-115-5 Sequence 5, Appl1
8 2182 96.6 1299 8 US-10-282-122A-7167 Sequence 7167, Ap
9 2182 96.6 1901 3 US-09-866-379-7 Sequence 7, Appl1
10 2182 96.6 1901 3 US-09-866-379-9 Sequence 9, Appl1
11 2182 96.6 1901 6 US-10-156-660-3 Sequence 3, Appl1
12 2182 96.6 1901 8 US-10-156-660-3 Sequence 7, Appl1
13 2182 96.6 1901 10 US-10-933-115-7 Sequence 7, Appl1
14 2182 96.6 1901 14 US-11-056-354-3 Sequence 3, Appl1
15 2181 96.6 1901 3 US-09-866-379-5 Sequence 5, Appl1
16 2175 96.3 1901 3 US-09-866-379-6 Sequence 6, Appl1
17 2175 96.3 1901 10 US-10-933-115-6 Sequence 6, Appl1
18 2171 96.1 1323 3 US-09-777-566A-1 Sequence 1, Appl1
19 2171 96.1 1323 3 US-09-866-379-1 Sequence 1, Appl1
20 2171 96.1 1323 6 US-10-034-985-1 Sequence 1, Appl1
21 2171 96.1 1323 7 US-10-430-356-1 Sequence 1, Appl1
22 2171 96.1 1323 8 US-10-601-319-1 Sequence 1, Appl1
23 2171 96.1 1323 10 US-10-933-115-1 Sequence 1, Appl1
24 2168.5 96.0 11357 13 US-11-074-522-14 Sequence 14, Appl1
25 2164 95.8 1486 13 US-11-018-709-2 Sequence 2, Appl1
26 2157 95.5 1281 7 US-10-334-672-4 Sequence 4, Appl1
27 2157 95.5 1281 7 US-10-334-671-4 Sequence 4, Appl1
28 2147 95.1 1489 6 US-10-266-041-9 Sequence 9, Appl1
29 2147 95.1 1489 7 US-10-284-962-1 Sequence 1, Appl1
30 2144 95.0 1486 7 US-10-284-962-4 Sequence 4, Appl1
31 2144 95.0 1486 13 US-11-018-709-4 Sequence 4, Appl1
32 1179 52.2 1737 10 US-10-450-763-25689 Sequence 25689, A
33 976.5 43.2 1326 6 US-10-021-723A-11 Sequence 11, Appl1
34 959.5 42.5 1326 8 US-10-282-122A-11608 Sequence 41608, A
35 954.5 42.3 1325 6 US-10-021-723A-3 Sequence 3, Appl1
36 934.5 41.4 1325 6 US-10-021-723A-1 Sequence 1, Appl1
37 782 34.6 1431 6 US-10-021-723A-9 Sequence 9, Appl1
38 723.5 32.0 4050 10 US-10-450-763-386 Sequence 386, App
39 713 31.6 466 8 US-10-282-122A-3181 Sequence 3181, Ap
40 637 28.2 1264 10 US-10-450-763-24247 Sequence 24247, A
41 637 28.2 1264 10 US-10-450-763-30120 Sequence 30120, A
42 554 24.5 11710 7 US-10-240-689-40 Sequence 40, Appl1
43 475.5 21.1 1230 6 US-10-021-723A-7 Sequence 7, Appl1
44 462.5 20.5 1266 6 US-10-021-723A-5 Sequence 5, Appl1
45 414 18.3 1326 16 US-11-036-686-9338 Sequence 9338, Ap

ALIGNMENTS

RESULT 1
US-11-056-354-1
; Sequence 1, Application US/11056354
; Publication No. US20050246780A1
; GENERAL INFORMATION:
; APPLICANT: SHORT, Jay M.
; APPLICANT: KREITZ, Keith
; APPLICANT: GRAY, Kevin A.
; APPLICANT: BARTON, Nelson R.
; APPLICANT: GARRETT, James B.
; APPLICANT: O'DONOGHUE, Eileen
; APPLICANT: MATUR, Eric J.
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING THEM AND METHODS
; TITLE OF INVENTION: FOR MAKING AND USING THEM
; FILE REFERENCE: 564462001803/D1370-9C1
; CURRENT APPLICATION NUMBER: US/11/056,354
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 10/156,660
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25

```

; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1296
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Modified phyase enzyme
US-11-056-354-1

Alignment Scores:
Pred. No.: 3,27e-255 Length: 1296
Score: 2258.00 Matches: 432
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 14 Gaps: 0

US-10-601-319-10 (1-432) x US-11-056-354-1 (1-1296)

QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
DB 1 ATGAAAGGAGTCTTAATCCCATTTTATCTCTTGATTCGGTTAAACCCCGCAATCTCA 60
QY 21 PheAlaGlnSerGluProGlnLeuLysLeuGlnSerValIleValSerArgHisGly 40
DB 61 TTCGTCGAGTGAAGCCGAGAGCTGAAGCTGGAAGATGTGGATGTCAGTGTGATGTT 120
QY 41 ValAlaGlnProThrLysAlaIleThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
DB 121 GTGCGTGTCCCAACCAAGGCCACGCACTGATGAGATGTCAACCCGACGACATGGCCA 180
QY 61 ThrTrpProValLysLeuGlnLysLeuThrProArgGlyGlnLeuIleAlaIleThrLeu 80
DB 181 ACCTGGCCGGTAAACCTGGGTGAGCTGAACCCGCGGTGGTGAACCTTAATGCCATATCTC 240
QY 81 GlyHisIleTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
DB 241 GCAACATTAATCGGCGTCAAGCTGTGTAGCCGACGATTTGCTCCCTAAATGTGGCTGCCG 300
QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGlnArgThrArgLysThrGlyGln 120
DB 301 CAGTCTGTGTCAGTGTGCGCATTTATGCTGATGTGACGAGCGTAAACCCGTTAAACAAGGCCAA 360
QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
DB 361 GCTTTCGCGCGGCTGCGACCTGACCTGTCGATCAATACCGTACATACCCAGGCGATACG 420
QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
DB 421 TCCAGTCCCGCATTCGTTATTTAATCTCTTAATAAATGCGGTTTGGCACTGGATTAACCG 480
QY 161 AsnValThrAspAlaIleLeuGlnArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
DB 481 AACGAGATCGACCGCATCTCGAAGAGGCGCAGAGGGGTCAATTTGCTGAATCCCGGCGAT 540
QY 181 TyrGlnThrAlaPheArgGlnLeuGlnArgValLeuAsnPheProGlnSerAsnLeuGly 200
DB 541 TATCAAAAGGGGCTTTCGGAACCTGSAACGGGGCTTAATTTTCGCCAATCAAACTGTGCG 600
QY 201 LeuLysArgGlnLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlnLeu 220
DB 601 CTTAAACCTGAGAAACAAGAGAAAGCTGTTCATTAAACGACGACATTCACATCGGAATCC 660
QY 221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
DB 661 AAGGTAGACGGCGCATGTGTCTCATTAACCGGTGCGTAAAGCTTCGATCATATGCTGACG 720
```

```

QY 241 GlnIlePheLeuLeuGlnGlnAlaGlnGlyMetProGlnProGlyTrpGlyArgIleThr 260
DB 721 GAGATATTTCTTCCTGCAACAGCAAGGAAATGCCAGACCGGAGTGGGAAGATCAACC 780
QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuGln 280
DB 781 GATTCAACACGATGGAACACTTGTCTAAGTTTGATTAACCGGCAATTTGATTTCTCAAA 840
QY 281 ArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
DB 841 CGCAGCCGACGAGGTGGCCCGGACGGCCACCCGTTATTGAATTCATCAACAGACGCG 900
QY 301 LeuThrProHisProProGlnLysGlnAlaIleArgValThrLeuProThrSerValLeu 320
DB 901 TTGAAGCCCATTCACCCGCAAAAACAGGCGTAATGTGTGATTAACCATTCCTGAGTCTG 960
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTrp 340
DB 961 TTATATCCGCGGACAGATACATCTGCAAAATCTCGCGCGCGCACCTGGAGCTCAACTGG 1020
QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlnLeuValPheGlnArgTrp 360
DB 1021 ACGCTTCCCGGTCAAGCGGATTAACACGCGCCAGGTGTGTGAATGTGTGAACGCTCG 1080
QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
DB 1081 CTTGGCTTAACCATTAACAGCAAGCATGATTCAGATTTCCGTGCTTCCAGACTTTACAG 1140
QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlnValLysLeuThr 400
DB 1141 CAGATGCGGTGTAATAACCGCGCTGTCAATTAAATACGCGCGCGGAGAGGTAAACTGACC 1200
QY 401 LeuAlaGlyCysGlnGlnArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
DB 1201 CTGGCAGATGTGAAGAGCAAAATGCCAGGCGATGTGTGTTGGTGAAGTTTAACCAA 1260
QY 421 IleValAsnGlnAlaArgIleProAlaCysSerLeu 432
DB 1261 ATCGTGAATGAACACGCAATACCGGCTGCAAGTTTG 1296

RESULT 2
US-10-156-660-1
; Sequence 1, Application US/10156660
; Publication No. US20030103958A1
; GENERAL INFORMATION:
; APPLICANT: Short, Day M.
; APPLICANT: Kretz, Keith
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson R.
; APPLICANT: Garrett, James B.
; APPLICANT: O'Donoghue, Eileen
; TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM
; TITLE OF INVENTION: AND METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 09010-029007
; CURRENT APPLICATION NUMBER: US/10/156,660
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
```

```
/ LENGTH: 1308
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: modified phyase enzyme
/ NAME/KEY: CDS
/ LOCATION: (1)...(1308)
US-10-156-660-1

Alignment Scores:
Pred. No.: 3,32e-255 Length: 1308
Score: 2258.00 Matches: 432
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-601-319-10 (1-432) x US-10-156-660-1 (1-1308)

QY 1 MetLyAlaIleuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
DB 1 ATGAAGAGCATCTTAATCCCATTTTATCTCTGATTCGGTTAACCCGCAATCTGA 60
QY 21 PheAlaGlnSerGluProGluLeuGlnSerValValIleValSerArgHisGly 40
DB 61 TTGCTCAGAGTGAGCCGAGCTGAAGCTGAAAGTGAGTGAATGTCAAGTCGTCAVGT 120
QY 41 ValAlaGlnProThrLyAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
DB 121 GTGGCGGTCCCAACCAAGCCCAACGCAATGATGACAGATCCCAACGATGCGCA 180
QY 61 ThrTrpProValIleLeuGlnIleuLeuThrProArgGlyGlyGlnLeuIleAlaTrpLeu 80
DB 181 ACCTGCGCCGCTAAACCTGGGTGAGCTGACACCGCGCGTGTGAGCTAATCCGCTATCTC 240
QY 81 GlyHisTrpTrpArgGlnArgLeuValAlaAspGlyLeuLeuProGlyCysPro 100
DB 241 GGCATTAATGCGGTGACGCTGAGGCGAGCGAGTCTGCTAAATGTGCTGCTCCG 300
QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGlnArgThrArgLyThrGlyGln 120
DB 301 CAGTCTGATCAGTGGCGCATTTATGCTGATGTCAGACGACGTAACCGTAAACAGCGCA 360
QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValIleThrGlnAlaAspThr 140
DB 361 GCCTTCGCGCGCGGCTGGCACCTGACTGTGCAATACGTAACCTACCGACGACATAG 420
QY 141 SerSerProAspProLeuPheAsnProLeuIleThrGlyValCysGlnLeuAspAsnAla 160
DB 421 TCCAGTCCCGATCCGTTATTTAATCTCTAAACCTGGGTTGCGCACTGATTAACGG 480
QY 161 AsnValIleThrAspAlaIleLeuGlnArgAlaGlySerIleAlaAspPheThrGlyHis 180
DB 481 AACGAGCATCGAGCGCATCTCCAGAGGCGAGAGGGTCAATGCTGACTTTACCGGGCAT 540
QY 181 TyrGlnIleThrAlaPheArgGlnLeuGlnArgValIleuAsnPheProGlnSerAsnLeu 200
DB 541 TATCAAAACGCGCTTTGCGCAACTGGAAACGGGTGCTTAATTTCCGCAATCAAACTTGTC 600
QY 201 LeuLyAspGlnIleGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlnLeu 220
DB 601 CTTAAACGAGAAACAGAGCAAGAGCTGTCTTAATTAACGACAGCATTAACATCGAAATCC 660
QY 221 LySValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
DB 661 AAGGTAGGCGCGAGCTGTCTCTTAACCGGTGCGGTAAAGCTCGCAATCAATGCTGACG 720
QY 241 GlnIlePheLeuLeuGlnAlaGlnAlaGlnIleMetProGlnProGlyTrpGlyArgIleThr 260
DB 721 GAGATATTTCTCTCTGACAAACAGACAGGAAATGCGGACCGGGGTGGGAAAGATCACCC 780
QY 261 AspSerHisGlnIleThrAsnThrLeuSerLeuHisAsnAlaGlnPheAspLeuGln 280
```

```
DB 781 GATTACACACAGTGAACACCTTGCTAAGTTGCAATACGCCCAATTTATTTGCTACAA 840
QY 281 ArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLyThrAla 300
DB 841 CCGACGCCAGAGAGTTGCCCGGAGCGCGCACCCCTTATTGATTTGATCAAGACAGCG 900
QY 301 LeuThrProHisProProGlnIleGlnAlaIleGlyValIleThrLeuProThrSerValLeu 320
DB 901 TTGACGCCCCATCCACCGCAAAACAGCGGTATGTGTGACATTAACCCACTTCAGTGTG 960
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValIleLeuIleuAsnTrp 340
DB 961 TTTATGCGCGGACACGATATCTAATCTGCAAAATCTCGGCGGCGCATGAGCTCAACTG 1020
QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlnLeuValPheGlnArgTrp 360
DB 1021 ACGCTTCCCGGTCAAGCGGATTAACAACCGCCAGAGTGTGAACTGTGTTGAACGCTGG 1080
QY 361 ArgAlaGlnSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
DB 1081 CCGTGGCTTAAGCGATTAACAGCAGTGAATTCAGGTTCCGCTGCTTCCAGACTTTACAG 1140
QY 381 GlnMetArgAspLyThrProLeuSerLeuAsnThrProProGlyGlnValIleLySerLeuThr 400
DB 1141 CAGATGCGGATTAACCGCGCTGTCTAATTAATACGCCCGCGAGAGGTGAACCTGACC 1200
QY 401 LeuAlaGlyCysGlnGlnArgAsnAlaGlnIleMetCysSerLeuAlaGlyPheThrGln 420
DB 1201 CTGGCAGATGTGAAGAGGAAATGCGCAGGCAATGTGTTCTGTGCAAGTTTACGCA 1260
QY 421 IleValAsnGlnAlaArgIleProAlaCysSerLeu 432
DB 1261 ATCGGAATGAAGACAGCATACCGCGCTGCACTTTG 1296

RESULT 3
US-10-601-319-9
/ Sequence 9, Application US/10601319
/ Publication No. US20040091968A1
/ GENERAL INFORMATION:
/ APPLICANT: Short, Jay M.
/ APPLICANT: Kretz, Keith A.
/ APPLICANT: Gray, Kevin A.
/ APPLICANT: Barton, Nelson Robert
/ APPLICANT: Garrett, James B.
/ APPLICANT: O' Donoghue, Eileen
/ APPLICANT: Mathur, Eric J.
/ TITLE OF INVENTION: RECOMBINANT PHYASES AND METHODS OF MAKING
/ TITLE OF INVENTION: AND USING THEM
/ FILE REFERENCE: 09010-029011
/ CURRENT APPLICATION NUMBER: US/10/601,319
/ PRIOR FILING DATE: 2003-06-20
/ PRIOR APPLICATION NUMBER: US 09/866,379
/ PRIOR FILING DATE: 2001-05-24
/ PRIOR APPLICATION NUMBER: US 09/580,515
/ PRIOR FILING DATE: 2000-05-25
/ PRIOR APPLICATION NUMBER: US 09/318,528
/ PRIOR FILING DATE: 1999-05-25
/ PRIOR APPLICATION NUMBER: US 09/291,931
/ PRIOR FILING DATE: 1999-04-13
/ PRIOR APPLICATION NUMBER: US 09/259,214
/ PRIOR FILING DATE: 1999-03-01
/ PRIOR APPLICATION NUMBER: US 08/910,798
/ PRIOR FILING DATE: 1997-08-13
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 9
/ LENGTH: 1308
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: modified phyase enzyme
```

```
! NAME/KEY: CDS
! LOCATION: (1)...(1296)
US-10-601-319-9

Alignment Scores:
Pred. No.: 3,32e-255 Length: 1308
Score: 2258.00 Matches: 432
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: Gaps: 0

US-10-601-319-10 (1-432) x US-10-601-319-9 (1-1308)

Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAAGGATCTTAATCCATTTTATCTCTTGATTCGATTACCCGCAATCTGCA 60
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValValIleValSerArgHisGly 40
Db 61 TTCGCTCAGAGTGAAGCCGAGACTGAAGCTGGAAGTGTGATGTCTCATCTCATGT 120
Qy 41 ValArgAlaProThrLysAlaIleThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 121 GTGGGTGCTCCAAACCAAGCCAGCACTGATGCAGATGTCACCCAGACAGCATGGCCA 180
Qy 61 ThrTrpProValLysLeuGlnLysLeuThrProArgGlyGlyLysLeuIleAlaTyrLeu 80
Db 181 ACCTGGCCGGTAAACTGGGTGAGCTGACCCGCGGTGGTGAAGTAAATGCTTATCTC 240
Qy 81 GlyHisTyrTrpArgGlnArgLysValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
Db 241 GGACATTATCTGGCGTCAACGCTGTGTACCGACGATTTGCTCCCTAAATGTGGCTGCCG 300
Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgLysThrGlyGln 120
Db 301 CAGTCTGCTCAGGCTCGCATTTATGCTGATGCTCCAGCCGATACCCGTAACAGCGCAA 360
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCTTCGCGCGGCTGCGACCTGACGTGTGAATACCGTACATACCCAGCGAGATACG 420
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCCAGTCCCGATCGCTTATTTATCTCTTAAAACTGGCGTTTGGCACTGGATACCG 480
Qy 161 AsnValThrAspAlaIleLeuGlnArgAlaGlyLysSerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGACTGACGCGATCTCTGAGAGGGCAGAGGGGTCAATTTGCTTAACCGGGCAT 540
Qy 181 TyrGlnThrAlaPheArgGluLeuGlnArgValLeuAsnPheProGlnSerAsnLeuGly 200
Db 541 TATCAAAAGGGGTTTCGAACTGGAACGGGGTCTTAATTTTCGCAATCAAACTGTGTC 600
Qy 201 LeuLysArgLysGlnLysAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlnLeu 220
Db 601 CTTAAACGTGAGAAACAGAGCAAGAGCTGTTCATTAAACGACGACATTTCATCGAATC 660
Qy 221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGACCGCGCATGTGTCTCACTTAACCGGTGGGTGAAGCTCTGCATCATGCTGACG 720
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnIleMetProGlnLysProGlyArgLysLeuThr 260
Db 721 GAGATATTCTCTGCAACAGCAAGGAAATGCCGGAGCGGGGTGGGAAAGATCAACC 780
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
Db 781 GATTCAACACGAGTGAACACCTGTAGTTTGCAATTAACGGGCATTTGATTGTACAA 840
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
Db 841 CGACGCGCAGAGGTTGCCGCGACGCGCCGACCCCGTTATTAGATTGATCAAGACAGCG 900
```

```
Qy 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 901 TTGAGCCCTCATTCACCGCAAAACAGCGGTATGTGTGACATTACCACTTCAGTGTCTG 960
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTrp 340
Db 961 TTATTCGCCGGAACAGATTAATCTGTGGCAAAATCTCGCGCGGCGACATGAGCTCACTGG 1020
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlyLeuValPheGluArgTrp 360
Db 1021 ACGCTTCCCGTCAACCGGATTAACACCGCCGACAGTGTGAATCTGTGTGAACGCTGG 1080
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1081 GTCGCTTAACGCAATTAACGCACTGATTCAGGTTTCCTGCTTCCAGACTTTACAG 1140
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlnValLysLeuThr 400
Db 1141 CAGATGCGTGAATAAACCGCGCTGTCACTTAAATACGCCGCCGAGAGGTGAACCTGACC 1200
Qy 401 LeuAlaGlyCysGluGlnArgAsnAlaGlnIleMetCysSerLeuAlaGlyPheThrGln 420
Db 1201 CTGGCAGAGATGTGAAGCGAAATGCGCAGGCGCATGTGCTTGGCAGGTTTACGCAA 1260
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1261 ATCGTAATGAAGCACCGCATACCGGCTGCAATTG 1296

RESULT 4
US-10-933-115-9
! Sequence 9, Application US/10933115
! Publication No. US20050281792A1
! GENERAL INFORMATION:
! APPLICANT: Short, Jay M.
! APPLICANT: Kretz, Keith A.
! APPLICANT: Gray, Kevin A.
! APPLICANT: Barton, Nelson Robert
! APPLICANT: Garrett, James B.
! APPLICANT: O' Donoghue, Eileen
! APPLICANT: Mathur, Eric J.
! TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES
! FILE OF INVENTION: THEREOF
! FILE REFERENCE: 564462001822
! CURRENT APPLICATION NUMBER: US/10/933, 115
! CURRENT FILING DATE: 2004-09-01
! PRIOR APPLICATION NUMBER: US/09/866, 379
! PRIOR FILING DATE: 2001-05-24
! PRIOR APPLICATION NUMBER: US 09/580, 515
! PRIOR FILING DATE: 2000-05-25
! PRIOR APPLICATION NUMBER: US 09/318, 528
! PRIOR FILING DATE: 1999-05-25
! PRIOR APPLICATION NUMBER: US 09/291, 931
! PRIOR FILING DATE: 1999-04-13
! PRIOR APPLICATION NUMBER: US 09/259, 214
! PRIOR FILING DATE: 1999-03-01
! PRIOR APPLICATION NUMBER: US 08/910, 798
! PRIOR FILING DATE: 1997-08-13
! NUMBER OF SEQ ID NOS: 10
! SOFTWARE: FastSeq for Windows Version 4.0
! SEQ ID NO 9
! LENGTH: 1308
! TYPE: DNA
! ORGANISM: Artificial Sequence
! FEATURE:
! OTHER INFORMATION: modified phytase enzyme
! FEATURE:
! NAME/KEY: CDS
! LOCATION: (1)...(1296)
US-10-933-115-9

Alignment Scores:
Pred. No.: 3,32e-255 Length: 1308
```



```

Qy      1 MetLysAlaIleuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
Db      188 ATGAAGAGGATCTTAATCCATTTTATCTCTTCGATTCGGTAAACCCGCAATCTCA 247
Qy      21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40
Db      248 TTCGTCGAGTGAAGCCGAGAGCTGAAGCTGGAAGATGTGTGATGTGATGTCATGCT 307
Qy      41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
Db      308 GTGCGTGTCCCAACCAAGGCCACGCACTGATGCGAAGTCAACCCGAGAGCAATGGCCA 367
Qy      61 ThrTrpProValLysLeuGlyGluLeuThrProArgGlyGlyLysLeuIleAlaTyrLeu 80
Db      368 ACCTGGCCGGTAAACTGGGTGGCTGACACCGCGNGGTGGTGAAGCTAATGCCCTATCTC 427
Qy      81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
Db      428 GGACATTAATGCGCGCACGCTCGGTAGCCGAGATTGCTGGCAAAAAGGGCTGCCCG 487
Qy      101 GlnSerGlyGlnValAlaIleIleLeuAspValAspGluArgThrArgLysThrGlu 120
Db      488 CAGTCTGTCAAGTTCGATTAATGCTGATGCGAGCGTAAACCCGTAAACAGCGCA 547
Qy      121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db      548 GCCTTCGCGCGCGGCTGGCACTGATCTGTGAAATACCGTAAATCCAGGAGATACG 607
Qy      141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db      608 TCACAGTCCCGAATCCCTTATTATCTCTAAATAATGGCGTTTGGCAATGGATTAACGG 667
Qy      161 AsnValThrAspAlaIleLeuGlnArgAlaGlyLysSerIleAlaAspPheThrGlyHis 180
Db      668 AACGAGCTGAGCGCATCTCAAGCGGAGGAGGCTCAATGCTGACCTTAACCGGCGAT 727
Qy      181 TyrGlnThrAlaPheArgGluLeuGlnArgValLeuAsnAspProGlnSerAsnLeuCys 200
Db      728 CGGCAACGGCGTTTCGGGAAGTGAACGGGTGCTTAATTTCCGCAATCAAACTTGTC 787
Qy      201 LeuLysArgGluLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlnLeu 220
Db      788 CTAAACCTGAGAAACAGAGCAAAAGCTTTCATTAAAGCGAGCATTAACATCGGAATCC 847
Qy      221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db      848 AAGGTGAGCGCGCAATGCTCATTAACCGGTGCGTGAAGCTGCAATGCTGACG 907
Qy      241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
Db      908 GAGATATTTCTCTCCCAACAAGCAAGCAAGGAATCCGAGCGCGGTGGGAAAGATCAC 967
Qy      261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuGln 280
Db      968 GATTACACACAGTGGAACACTTGCTAAGTTTGCATTAACGCGCAATTTATTTGCTACAA 1027
Qy      281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleThrAla 300
Db      1028 CGCACGCAAGAGTTGCCGCGAGCGCGCAACCCGTTAATTAAGTTTAACAAGCACG 1087
Qy      301 LeuThrProHisAspProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db      1088 TTGACGCCCATCCACCCCAAAAACAGCGGTATGTGTGACATTCACCACTTCAGTGTG 1147
Qy      321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTrp 340
Db      1148 TTTATCGCGGACAGATTAATCTGCAATATCTGCGCGGCGCACTGAGGCTCAACTGG 1207
Qy      341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGlnArgTrp 360
Db      1208 ACGCTTCCCGGTACCCGGATTAACAGCGCGGAGGTGTGAACGTGTGTGAAACGCTGG 1267

```

```

Qy      361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db      1268 CGTGGCTAAGCAGTAACAGCAAGTTCAGTTTCGTCGTTCCTCCAGACTTTACAG 1327
Qy      381 GlnMetLysAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
Db      1328 CAGATGGGTGATTAACAGCCGCTGTCAATTAATCGCGCGGAGAGGTGAACCTGACC 1387
Qy      401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db      1388 CTGGCAGAGATGTGAAGACCAAAATGCCAGGCGCATGTGTTGCTTGGCAGGTTTAACGCA 1447
Qy      421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db      1448 ATCGTGAATGAACACCGCATACCGCGCTGCAGTTTG 1483

RESULT 6
US-10-601-319-5
; Sequence 5, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Machur, Eric J.
; TITLE OF INVENTION: RECOMBINANT PHRYASES AND METHODS OF MAKING
; TITLE OF INVENTION: AND USING THEM
; FILE REFERENCE: 09010-029011
; CURRENT APPLICATION NUMBER: US/10/601, 319
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 403
; OTHER INFORMATION: n = A,T,C or G
US-10-601-319-5

Alignment Scores:
Pred. No.: 5,94e-247 Length: 1901
Score: 2190.00 Matches: 425
Percent Similarity: 98.4% Conserves: 0
Best Local Similarity: 98.4% Mismatches: 7
Query Match: 97.0% Indels: 0
DB: 8 Gaps: 0

US-10-601-319-10 (1-432) x US-10-601-319-5 (1-1901)
Qy      1 MetLysAlaIleuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
Db      188 ATGAAGAGGATCTTAATCCATTTTATCTCTTCGATTCGGTAAACCCGCAATCTCA 247
Qy      21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40

```

```

Db      248 TTCGCTCAGAGTGAAGCCGAGCTGAAGCTGAAGAGTGTGATTTGTCAGTGTCAATGCT 307
Qy      41  VLLAAGALAPROTHRYALATHRGILNLEUMETGINASPVALTNRPROASPALATRP 60
Db      308 GTGGGTCTCCAAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 367
Qy      61  THTTPROVALYLYLEUGLYLEUETHRPROARGLYGLYGLULEUTHLEALATYRLEU 80
Db      368 ACCTGGCCCGTAAACTGGGTGAGTGAACCGCGAGGTGTGAGTGAATTCGCTATCTC 427
Qy      81  GLYHISTYTPARGLINARGLEUVALAASPGLYLEUETHRPROLYSCYSGLYCYVPRO 100
Db      428 GGAACATTAACCAAGCCAGCGCTCTGGTAGCGACGAGTTCTGGCGAAAAAGGCGTCCGG 487
Qy      101 GINSERGYGLINVALAALILEILEALASPVALLASPGILNARGLYSTRGYSTRGYLU 120
Db      488 CAGTCTGGTCAGGTGCGGATTAATGTGTGATGTGACGACGACGACGACGACGACGACG 547
Qy      121 ALAPHEALALAGLYLEUALAPROASPGYVALAILETHRVALLHISTHRGILNLAASPTHR 140
Db      548 GCCTTCGCGCGCGCGGCTGGCACTTGACTGTGCAATTAACCGTACATACCAAGGAGATACG 607
Qy      141 SERSERPROASPGPROLEUENHASNPROLEULYSTRGYVALCYSGILNLEUASPHANALA 160
Db      608 TCCAGTCCCGATCCGTTATTTAATCCTTAAAAAAGTGGGTTTGGCAACTGAGATACGCG 667
Qy      161 AENVALLTHRAPALALILEUGLUARGALAGLYGYSERILEALASPHETHRGILYHIS 180
Db      668 AACGAGACTGACGACATCTCTCAGCAGGCGAGGAGGCTCAATTGACTTTCACGCGGCA 727
Qy      181 TYRGINTHRALAPHEARGILEUGILNARGVALLEUASPHPROGLISERSENLEUCYS 200
Db      728 CGGCAAGCGCGCTTTCGCAACTGGAACGGGTGCTTAATTTTCGCAATCAACCTTGTC 787
Qy      201 LEULYARGVGLULYGLINASPGILUSERCYSSERLEUTHRGILNLALEUPROSERGLUEN 220
Db      788 CTTAAACGTGAGAAAACAGGACGAAGCTGTTCAATTAACGACGACATTAACATCGAATCC 847
Qy      221 LYSEVALSERLAAASPCYVALSERLEUTHRGILYVALSERLEUALASERLEUTHR 240
Db      848 AAGGTGAGCGCGCAATGCTCTCAATTAACCGGTGCGGTGACCTGCATCAATGCTGACG 907
Qy      241 GIULIPELEULEUGILNLAAGILNLYMETPROGLUPROGLYTRPGILYARGILETHR 260
Db      908 CAGATATTTCTCTGCAACACGACGAGGAAATGCGGAGCGCGGAGGAGGAGATCC 967
Qy      261 AASPERHLSGINTRPASNTHRLEUENUSERLEUHLASNALAGLNPHEASPLEUENGIN 280
Db      968 GATTCACACCAAGTGAAACCTTGTAAGTTGGCATTAACGCGCATTTTATTTGCTACA 1027
Qy      281 ARGTHRPROGLIVALAALASGSEARVALATHRPROLEULEUASPLEULILEYSTRALA 300
Db      1028 CGCAGCGCAAGAGTTGCCGCGACGCGCGCACCCGTTATTAAGATTGATCAAGACAGCG 1087
Qy      301 LEUTHRPROHLSAPROPROGLINLYSGILNATYRGLYVALTHLEUPROTHSERFALLEU 320
Db      1088 TTGAGCGCCCATCCACCGCAAAAACAGGCGTATGTGTGACCTTAACCACTTCACTGCTG 1147
Qy      321 PHEILEALAGLYHLSASPTHRASNLEUALAASLEUGLYGLYVALALEUGILULEUASNTR 340
Db      1148 TTTATCGCGGACAGATATTAATCTGGCAATCTCGGCGGCGCACTGAGTCAACTCG 1207
Qy      341 THRLEUPROGLYGLINPROASPHANTHRPROPROGLYGLYGLULEUVALPHEGLNARGTRP 360
Db      1208 ACGCTTCCCGGTCAACCGGATTAACAGCGCGCGAGGTGTGAACTGTGTTGMAAGCTGCG 1267
Qy      361 ARGAGYLEUSERASPSERGLINTRPILEGLINVALSERLEUVALPHEGLNTHREUGIN 380
Db      1268 CGTCGGCTTAAGCATTAACGACGATTAACGATTAACGATTAACGATTAACGATTAACG 1327
Qy      381 GINWELCARASPLYSTRPROLEUSERLEUANTHRPROPROGLYGLYVALYLYLEUTHR 400
Db      1328 CAGATGCTGATTAACGCGCGCTGTCAATTAATAGCGCGCGCGGAGAGGTGAATGAC 1387

```

```

Qy      401 LEUALAGLYCYSGILNARGVALAAGILNLYMETCYSSERLEUALAGLYPHETHRGIN 420
Db      1388 CTGGCAGATGTGAAGAGGAAATCGCAGGCGCATGTGTTCTTGGCAGGTTTTCGCAA 1447
Qy      421 ILEVALASNGUALAARGILEPROALAASPSERLEU 432
Db      1448 ATCGTGAATGAAGACGACATACCGGCGGTGCAATTTG 1483

RESULT 7
US-10-933-115-5
; Sequence 5. Application US/10933115
; Publication No. US20050281792A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES
; FILE REFERENCE: 5644620182
; CURRENT APPLICATION NUMBER: US/10/933,115
; PRIOR FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: 403
; OTHER INFORMATION: n = A,T,C or G
US-10-933-115-5

Alignment Scores:
Pred. No.: 5,94e-247 Length: 1901
Score: 2190.00 Matches: 425
Percent Similarity: 98.4% Conservative: 0
Best Local Similarity: 98.4% Mismatches: 7
Query Match: 97.0% Indels: 0
Gaps: 0

US-10-601-319-10 (1-432) x US-10-933-115-5 (1-1901)
Qy      1  MELYSALALELEULEIPROPHLEUSERLEULEULEIPROLEUTHRPROGLINSERALA 20
Db      188 ATGAAAGGATCTTAATCCATTTTATCTTGTGATTCGTTAAACCCGCAATCTGCA 247
Qy      21  PHEALGINSERGLUPROGLINLEUETHRLEUENUSERVALIILEVALSERARGHISGLY 40
Db      248 TTCGCTCAGAGTGAAGCCGAGACTGAAGCTGAAGAGTGTGATTCATGCTGATGCT 307
Qy      41  VALRGAAPROTHRYALATHRGILNLEUMETGINASPVALTNRPROASPALATRP 60
Db      308 GTGGGTCTCCAAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 367
Qy      61  THTTPROVALYLYLEUGLYLEUETHRPROARGLYGLYGLULEUTHLEALATYRLEU 80

```

```
Db 368 ACCGCGCCGCTAAAGTGGTGAAGTCAACCGCGGCTGGTGAAGTCAATGCGCTATCTC 427
Qy 81 GYHISYTRTPAAGLINAAGLNUVAIAAASPGLYLEUENUPOLYSCYEGLYCYEPRO 100
Db 428 GGAATTAACCAACGCGCGTCTGGTACCCACGATTTGCTGGCAAAAAGGCGTCCCG 487
Qy 101 GINSERGYGINVAIAAIEIAAASPVALASPGLUAAGTTPARXGLYETHGLVGLU 120
Db 488 CAGCTCTGATCAGGTGCGCATTTATGCTGATCTCAGCGAGCTTACCCGTAAACAGGCGAA 547
Qy 121 AAlphealaaaglyeuuaiproaspysaiaaiaethraiaisthrginaaaspthr 140
Db 548 GCTTCGCGCGCGGCTGGCACTGACTGCAATTAACCGATACATACCAGGAGATACG 607
Qy 141 SerserProaspProleuPhaenProleuylsthriglyvalCYEGINleuaspasna 160
Db 608 TCCAGTCCCGATCCGTTATTTAAATCTCTTAAAACTGGCGTTTGGCAATGATTAACGCG 667
Qy 161 AsnvalThraapAaiaieueugluarvalaglyseriaaiaaspPhethrglyhis 180
Db 668 AACGTGATTAACGCGATCTTCAAGGAGGAGGAGGCTCAATTTGCTGACTTAAACGCGCAT 727
Qy 181 TyrGINthraalapeargluleugluarvalleuaspheproginserasnuCYs 200
Db 728 CGGCAACGCGGCTTCCGGAACGTGGAACGGGTCTTAATTTCCGCAATCAACTGTGTC 787
Qy 201 LeulysaaglyulysgluaspgluserCYserleuthrginaaleuProsergluleu 220
Db 788 CTTAAACCTGAGAAACAGAGACGAAGCTGTTCAATTAAACGCGCAATTCATCCGAACTTC 847
Qy 221 LysValseraiaaspCYvalserleuthrglyalValserleuaiaSerMerleuthr 240
Db 848 AAGGTGACCGCGCAATGCTCTTAACCGGTGCGTAAGCTTGATCAATGCTGACG 907
Qy 241 GluilepheleuenglnaiaaglnlymecProgiupProgiYTRGlyArgllethr 260
Db 908 GAGATATTTCTCTGCAACAAGCAACAGGAAATGCCGAGCGCGGTGGGAAGATCAACC 967
Qy 261 AspserhisgINTTPAsnthrleuSerleuHisanaaiaaglnPheaspLeuengln 280
Db 968 GATTCAACACCAAGTGAAACACTTGTCTAAGTTGCAATTAACGCGCAATTTATTTGTCAAA 1027
Qy 281 ArgThrProgiuValaIarSeraarvalaThrProleuLeuAspLeuilesthrzla 300
Db 1028 CGCAGCGCAAGGTTGCCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1087
Qy 301 LeuthrProhisaproProginlysglnaiaarglyvalThrleuProThrservalieu 320
Db 1088 TTGACGCCCCCATCCACCGCAAAAACAGGCGGTATGTGATGACATTAACCACTCACTGCTG 1147
Qy 321 PheilealaglyhisAAspThrasleuuaiaaenleuglyvalleuengluenaanttrp 340
Db 1148 TTTATCGCGGACAGAGATTAATCTGCAATCTGCGCGGCGGCACTGAGGTCAACATCGG 1207
Qy 341 ThrleuProgiYGLINProAspAsnthrProProgiYGLYGLUleuValPhegluargtrp 360
Db 1208 ACGCTTCCCGGTCAAGCCGATTAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1267
Qy 361 ArgArgleuSerAspAsnsergINTTPIleginaaSerleuValPheGINthrlengln 380
Db 1268 COTGGCTTAAGGATTAACAGCAGCATGATTCAGTTTGGCTGCTTCCAGACTTTAACG 1327
Qy 381 GlnMetAagAspLythrProleuSerleuAsnthrProProgiYGLUVallyleuThr 400
Db 1328 CAGAGCGGTGATTAACCGCGCTGTCTATTAATAGCGCGCGCGGAGAGGTGAACATGACC 1387
Qy 401 LeuAaaglyCYsgluarvalasnaaiaaglnlymetCYserleuAaaglyPheThrGIN 420
Db 1388 CTGGGAGATGTGAAGAGGAAATGCGAGGCGATGTGTCGTGGCAAGGTTTACGCA 1447
Qy 421 IleValaengluAaarglleProaiaCYserleu 432
```

```
Db 1448 ATCGTGAATGAAAGACGACATACCGCGCTGCAGTTTG 1483
RESULT 8
US-10-282-122A-7167
; Sequence 7167, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyekind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7167
; LENGTH: 1299
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-282-122A-7167
Alignment Scores:
Pred. No.: 3,03e-246 Length: 1299
Score: 2182.00 Matches: 424
Percent Similarity: 98.1% Conservative: 0
Best Local Similarity: 98.1% Mismatches: 8
Query Match: 96.6% Indels: 0
Gaps: 0
DB:
US-10-601-319-10 (1-432) x US-10-282-122A-7167 (1-1299)
Qy 1 MetLysAaiaieueuileProPhaLeuSerleuileProleuThrProginseraia 20
Db 1 ATGAAAGGATCTTAATCCATTTTATCTTCTGATTCGCTTAACCCCGCAATCTGCA 60
Qy 21 PheAaagInsergluProgluleuylsleugluserValIleValSerArgHisgLY 40
Db 61 TTCGTCAAGATGAGCGCGGAGCTGAAGCTGAAGTGGTGGATGTGATGTCATGCTGT 120
Qy 41 ValArgAlaProthrLYsValaThrgInleuMetGlnaPvalThrProaspAlaTrpPro 60
Db 121 GTGGGTGCTCCAAACCAAGCGCACGCAACTGATGACGATGTCAACCCAGAGCATGGCCA 180
```

QY 61 ThrTPProValIysLeuGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
 DB 181 ACCTGGCCGGTAAACTGGGTGGCTGACACCGCGGAGTGAGCTAATCGCTAATCTC 240
 QY 81 GlyHisTyrTPArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
 DB 241 GGACATTAACCAACGCGCAGCGCTGGGTAGCCGAGTTGCTGGCGAAAAGGGCTGCCG 300
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 DB 301 CAGTCTGGTCAGAGTGGCGATTATGCTGATGTCGACGACGTCATCCGTTAAACAGCGCA 360
 QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnIleAspThr 140
 DB 361 GCCTTCGCGCGCGGCTGGCACCCTGACCTGCAATACCGTACATACCCAGCGCAAGTACG 420
 QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 DB 421 TCCAGTCCCGATCCGTTATTTAATCCTTAAACACGCGGTTGGCACTGGATTAACGCG 480
 QY 161 AsnValThrAspAlaIleLeuGluArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 DB 481 AACGTGACTGACGCGCATCTCTACAGCGGCGAGGAGGCTCAATTGCTGACTTTACCGGCGCAT 540
 QY 181 TyrGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
 DB 541 CGGCAACGCGCGCTTTCGCAACTGGAAACGGGTGCTTAATTTCCGAACTAACTTGTC 600
 QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 DB 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTACGACGATTAACATCGGAACCTC 660
 QY 221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 DB 661 AAGGTGACCGCGCAATGTCCTCATTTAACCGGTGCGGTAGCCTGCATCAATGCTGACG 720
 QY 241 GluIlePheLeuLeuGlnGlnIleAlaGlnGlyMetProGlnProGlyThrGlyArgIleThr 260
 DB 721 GAGATATTTCTCTCCAAACAGCACAGGAAATCCGAGCCGCGGTGGGAAAGGATCACCC 780
 QY 261 AspSerHisGlnThrAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
 DB 781 GATTACACACGATGGAACACCTTGCTAAGTTGCTAAGCCGCAATTTATTTGCTACAA 840
 QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
 DB 841 CGCAGCGCAAGAGGTGGCCGCGCAGCGCGCACCCGTTATGATTTGATCAAGACAGCG 900
 QY 301 LeuThrProHisLeuProGlnLysGlnIleAlaTyrGlyValThrLeuProThrSerValLeu 320
 DB 901 TTGAGCGCCCATCCACCGCAAAACAGGAGTATGAGTGAACATTACCCACTTCACTGCTG 960
 QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnThr 340
 DB 961 TTTATCGCGGACACGATTAATCTGCAATCTCGCGGCGCGCACTGAGCTCAACTCG 1020
 QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTyr 360
 DB 1021 ACGCTTCCCGGTCAACCGGATTAACACGCGCGCAGAGTGTGAACCTGTTTGAACGCTGG 1080
 QY 361 ArgArgLeuSerAspAsnSerGlnThrIleGlnValSerLeuValPheGlnThrLeuGln 380
 DB 1081 CGTCCGCTTAAGGATTAACGCGAGTGAATTCAGGTTTGGCTGGCTTTCCAGACTTTACG 1140
 QY 381 GlnMetCysAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
 DB 1141 CAGATCGGTGATTAACCGCGCTGTCTTAATTAACGCGCGCGGAGAGGTGAACCTGACC 1200
 QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 DB 1201 CTGGCAGATGTGAAGAGGAAATGCGCAGGCGCATGTGTGCTTGGCAGAGGTTTACGCA 1260

QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
 DB 1261 ATCGTGAATGAAGACGCAATACCGCGGTGCACTTTG 1296
 RESULT 9
 US-09-866-379-7
 ; Sequence 7, Application US/09866379
 ; Patent No. US20020136754A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DIVERSA CORPORATION
 ; APPLICANT: SHORT, Jay
 ; APPLICANT: KRETZ, Keith
 ; APPLICANT: GRAY, Kevin
 ; APPLICANT: BARTON, Nelson
 ; APPLICANT: GARRETT, James
 ; APPLICANT: O'DONOGHUE, Eileen
 ; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
 ; FILE REFERENCE: DIVER1370-7
 ; CURRENT APPLICATION NUMBER: US/09/866,379
 ; PRIORITY FILING DATE: 2001-05-24
 ; PRIORITY APPLICATION NUMBER: US 09/580,515
 ; PRIORITY FILING DATE: 2000-05-25
 ; PRIORITY APPLICATION NUMBER: US 09/318,528
 ; PRIORITY FILING DATE: 1999-05-25
 ; PRIORITY APPLICATION NUMBER: US 09/291,931
 ; PRIORITY FILING DATE: 1999-04-13
 ; PRIORITY APPLICATION NUMBER: US 09/259,214
 ; PRIORITY FILING DATE: 1999-03-01
 ; PRIORITY APPLICATION NUMBER: US 08/910,798
 ; PRIORITY FILING DATE: 1997-08-13
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 7
 ; LENGTH: 1901
 ; TYPE: DNA
 ; ORGANISM: Escherichia coli
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)-(1901)
 ; OTHER INFORMATION: n is any nucleotide
 US-09-866-379-7
 Alignment Scores:
 Pred. No.: 5,226-246 Length: 1901
 Score: 2182.00 Matches: 424
 Percent Similarity: 98.1% Conservative: 0
 Best Local Similarity: 98.1% Mismatches: 8
 Query Match: 96.6% Indels: 0
 DB: 3 Gaps: 0
 US-10-601-319-10 (1-432) x US-09-866-379-7 (1-1901)
 QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 DB 188 ATGAAGAGGATTTATATCCATTTTATCTCTTCGATTCGGTTAACCCCGCAATCTGCA 247
 QY 21 PheAlaGlnSerGluProGluLeuLysLeuSerValIleAlaSerArgHisGly 40
 DB 248 TTCGCTCAGAGTGAAGCCGGAAGCTGAAGCTGAAGAGTGTGATGTCAGTGCATGCT 307
 QY 41 ValArgAlaProThrIleValAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
 DB 308 GTGCGTGTCCAAACGCAAGCCGCAAGCTGATGACGAGATGTCACCCCAAGCGCATGGCCA 367
 QY 61 ThrTPProValIysLeuGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
 DB 368 ACCTGGCCGGTAAACTGGGTGGCTGACACCGCGGAGTGTGAGCTAATGCGCTAATCTC 427
 QY 81 GlyHisTyrTPArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
 DB 428 GGACATTAACCAACGCGCAGCGCTGGGTAGCCGACGATTTGCTGGCGAAAAGGGCTGCCG 487
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120

Db 488 CAGCTGTCAGCTGCGCATTTATCTCATATGTCACGACGACGATCCCGTAAACGAGCGAA 547
Qy 121 AlAphaealAaAglyLeuAlaProAspCySaAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCCTTCGCGCGCGGCTGCGACCTGACGTGCAATTAACCGATACATACCGACGAGATACG 607
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 608 TCCAGTCCCGCATCGGTATTTAAATCTCTTAAATAATGCGGTTCGCAACTGATTAACGCG 667
Qy 161 AsnValThrAspAlaIleLeuGlnArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db 668 AACGTGACTGACGCGATCTCTACGACGCGCAGAGGCTCAATTGCTGATTCATCCGGCAT 727
Qy 181 TyrGlnThrAlaPheArgGlnLeuGlnArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 728 CCGGAAACGCGGTTTCGCAACTGGAACTGGAACTGATTTATTTCCGCAATCAACTTGTC 787
Qy 201 LeuLysArgGlnLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlnLeu 220
Db 788 CTTAAACGTGAAGAACAGACGAAAGCTGTTCAATTAAACGACGACATTAACGGAATC 847
Qy 221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 848 AAGGTGACCGCGCAATGTCATTAACCGGTGCGTAAAGCTTCGATCAATGCTGACG 907
Qy 241 GlnIlePheLeuLeuGlnGlnAlaGlnSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 260
Db 908 GAGATATTTCTTCGCAACAGACACAGGAAATGCGGAGCGCGGATGAGGAAAGATCAC 967
Qy 261 AspSerHisGlnTPAsnThrLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
Db 968 GATTTCACACGAGTGAACACCTTGCTAAGTTTGCAATACGGCAATTTATTTGCTACAA 1027
Qy 281 ArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
Db 1028 CGCAGCGCAGAGGTTGCCCGACGCGCCACCCCGTTATTAAGTTTATGATCAAGACGCG 1087
Qy 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 1088 TTGACGCGCCCATCCACCGCAAAAACAGCGGATGATGATTAACCATTCACAGTGCTG 1147
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnThr 340
Db 1148 TTATTCGCGGACACGATATCAATCTGGCAAAATCTGGCGGCGCACCTGACCTCAACTG 1207
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlnLeuValPheGlnArgThr 360
Db 1208 ACGCTTCCCGGTCAAGCCGATTAACGCGCCGCAAGTGTGAACGTGTTGAAACGCTG 1267
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1268 CGTGCGCTTAACCGATTAACAGCACAGTGATTCAGGTTTCGCTGCTTCACAGCTTAA 1327
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
Db 1328 CAGATGCGGTAAACCGCCCTGTCAATTAATACGCGCCCGGAGAGATGAAACCTGACC 1387
Qy 401 LeuAlaGlyCysGlnGlnArgAsnAlaGlnIleLysMetCysSerLeuAlaGlyPheThrGln 420
Db 1388 CTGGCAGAGATGTAAGAGCGAAATCCGACGAGGCAATGTTTCGTTGACGAGTTTAA 1447
Qy 421 IleValAsnGlnAlaArgIleProAlaCysSerLeu 432
Db 1448 ATCGTGAATGAAGCACGCAATACCGCGCTGACAGTTTG 1483

RESULT 10
US-09-866-379-9
; Sequence 9, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION

APPLICANT: SHORT, Jay
APPLICANT: KRETZ, Keith
APPLICANT: GRAY, Kevin
APPLICANT: BARTON, Nelson
APPLICANT: GARRETT, James
APPLICANT: O'DONOGHIE, Bileen
TITLE OF INVENTION: RECOMBINANT BACTERIAL PHAYASES AND USES THEREOF
FILE REFERENCE: DIVERSA370-7
CURRENT APPLICATION NUMBER: US/09/866,379
CURRENT FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: US 09/580,515
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 09/318,528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: US 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 09/259,214
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: US 08/910,798
PRIOR FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.1
SEQ ID NO 9
LENGTH: 1901
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(1901)
OTHER INFORMATION: n is any nucleotide
US-09-866-379-9
Alignment Scores:
Pred. No.: 5.22e-246 Length: 1901
Score: 2182.00 Matches: 424
Percent Similarity: 98.1% Conservative: 0
Best Local Similarity: 98.1% Mismatches: 8
Query Match: 96.6% Indels: 0
Gaps: 0
US-10-601-319-10 (1-432) x US-09-866-379-9 (1-1901)
Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 188 ATGAAGCGATCTTAATCCATTTTATCTCTTGAATTCGCTTAACCCGCAATCTGCA 247
Qy 21 PheAlaGlnSerGlnProGlnLeuLysLeuGlnSerValValIleValSerArgHisGly 40
Db 248 TTGCTCAGAGTGAAGCCGAGCTGAAGCTGGAAGTGTGGATTCAGTCGTCAATGCT 307
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 308 GTGCTGCTCCAAACCAAGGCGCACGCACTGATGAGATGTCAACCCAGACGATGCCA 367
Qy 61 ThrTrpProValLysLeuGlyGlnLeuThrProArgIleGlyGlnLeuIleAlaTyrLeu 80
Db 368 ACTTGCGCGTAAACCTGGGTGCTGACACCGCGAGTGTGAAGTGAAGTGAAGTGAAGT 427
Qy 81 GlyHisTyrTrpArgGlnArgLeuValAlaAspIleLeuLeuProLysCysGlyCysPro 100
Db 428 GGAATTATCAACGCGCGCTGCTGAGCCGAGATTCGCTGGGGAAGGAGGCTGCCG 487
Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGlnArgThrArgLysThrGlyGln 120
Db 488 CAGCTGCTGACGCTGCGGATTTATGCTGATGTCACAGAGGTACCCGTTAAACGAGGAA 547
Qy 121 AlAphaealAaAglyLeuAlaProAspCySaAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCCTTCGCGCGCGGCTGCGACCTGACTGTGCAATTAACCGATACATACCGACGAGATACG 607
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 608 TCCAGTCCCGCATCGGTATTTAAATCTCTTAAATAATGCGGTTCGCAACTGATTAACGCG 667

```

QY 161 AsnValThrAspAlaIleuGluArgAlaGlySerIleAlaAspPheThrGlyHis 180
DB 668 AACGAGACGACGCGCATCTCTACAGCGGCGAGAGGGTCAATTCCTACCTTACCGGGCAT 727
QY 181 TyrGlnThrAlaPheArgGluLeuGluArgValIleuAsnPheProGlnSerAsnLeuCys 200
DB 728 CCGCAAAACCGCTTCCGCACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTC 787
QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
DB 788 CTTAAACGTGAAGAAACAGACGAAACCTGTTCAATTAACGAGCATTAACATCGGAATC 847
QY 221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerLeuThr 240
DB 848 AAGGAGGCGCGGACCAATCTCTATTAACCGGTGCGGTAAAGCTCCGCAATCAATGCTGAG 907
QY 241 GlnIlePheLeuLeuGlnIleAlaGlnIleMetProGluProGlyTyrGlyArgIleThr 260
DB 908 GAGATATTTCTCTGCAACAGACAGGGAATGCCGAGCCGGGTGGGGAAGATCAAC 967
QY 261 AspSerHisGlnIleThrAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuGln 280
DB 968 GATTCAACACAGTGAACACCTTGCTAAGTTGCAATMACGCCAATTTGTTGCTACAA 1027
QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
DB 1028 CGCAGCGCCAGAGTTGCCGCGACCGCGCCGCGCTATTAATTAATTAACAGACAGCG 1087
QY 301 LeuThrProHisArgProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
DB 1088 TTGACGCCCATCCACCGCAAAACAGGGGTATGGGTACATTAACCATTCAGTGGCG 1147
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeuGluLeuAsnTrp 340
DB 1148 TTTATCGCGGACAGCATATCAATCTGCAATCTCGGGGCGCACTGAGCTCAACTGG 1207
QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGlnArgTrp 360
DB 1208 ACGCTTCCCGGTACGCGCATACACGCCGCGAGGTGTGAACGTGGTTTAAACGCTGG 1267
QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
DB 1268 CGTGGCTTAAGCATTAACAGCGAGTTCAGGTTTGGCTGCTTCCAGACTTTACAG 1327
QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
DB 1328 CAGATGCGGTAAACCGCGCTGTCAATTAATACGCGCGCGGAGAGGTGAACGTAC 1387
QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnIleMetCysSerLeuAlaGlyPheThrGln 420
DB 1388 CTGGCAGATGTGAAGCGGAATCGCAGGCGCATGTGTCTGTGGCAGGTTTAAACGCA 1447
QY 421 IleValaGlnAlaArgIleProAlaCysSerLeu 432
DB 1448 ATCGGATGAAGCAGCATACCGCGCTGCACTTTG 1483

```

```

; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (188) ... (1483)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 403
; OTHER INFORMATION: n = A,T,C or G
US-10-156-660-3
US-10-601-319-10 (1-432) x US-10-156-660-3 (1-1901)
Alignment Scores:
Pred. No.: 5,226-246 Length: 1901
Score: 2182.00 Matches: 424
Percent Similarity: 98.1% Mismatches: 0
Best Local Similarity: 98.1% Indels: 0
Query Match: 96.6% Gaps: 0
DB: 6
QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
DB 188 ATGAAGGAGATCTTAATCCATTTTATCTCTTCAATCCGTTAAACCCGCAATCTGCA 247
QY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40
DB 248 TTCGCTCAGAGTGAGCCGAGCTGAAGCTGAAGTGGTGAATGTCAAGTGTGATGCT 307
QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
DB 308 GTGCGTGTCCACCAACCAAGCCAGCACTGATGCAAGATGCACCCAGAGCATGGCCA 367
QY 61 ThrTrpProValLysLeuGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
DB 368 ACTGCGCGGTAAAACTGGGTGGCTGACACCGCGGGTGGTGAAGTCAATCGCTATCTC 427
QY 81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
DB 428 GACATTAACCAACGCGCAAGCTGTGTAGCCAGCATTTGCTGGCGGAAAAAGGCTGCCG 487
QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGly 120
DB 488 CAGTCTGTCAAGTCTCGCATTTTCTGATGTCAACAGGCTTACCCGTAAACAGGCGAA 547
QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
DB 548 GCTTTCGCGCGCGGCTGGCAGCTGACGTGCAATTAACCGTACATCAACCAAGAGATACG 607
QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
DB 608 TCCAGTCCCGATCGGTATTTAATCTCTTAAAACTGGCGTTTCCAACTGATTAACCGCG 667
QY 161 AsnValThrAspAlaIleLeuGluArgAlaGlySerIleAlaAspPheThrGlyHis 180
DB 668 AACGAGACGACGCGCATCTCTACAGCGGCGAGAGGGTCAATTCCTACCTTACCGGGCAT 727

```

```
QY 181 TyrGlnThrAlaPheArgGluLeuGlnuArgValIleuAenPheProGlnSerAsnLeuCy 200
Db 728 CGGCAAAAGCGGTTTCGGAACTGGAACGGGAGTCTTAATTTCCGAATCAAACTGTGCG 787
QY 201 LeuYsArgGlnYsGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGlnu 220
Db 788 CTTAAACGTGAGAAACAGAGCAAGAAAGCTGTTCACTTAACGACGACTTCCATCGGAATC 847
QY 221 LysValSerAlaAspCysValSerLeuThrGlnAlaValSerLeuAlaSerMetLeuThr 240
Db 848 AAGGTGAGCGCCGCAATGTCCTCACTTAACCGGTGGTGAACCTCGCATCAATGCTGACG 907
QY 241 GlnIlePheLeuLeuGlnGlnAlaGlnIleMetProGluProGlnYTrpGlnArgIleLeu 260
Db 908 GAGATATTCTCTCGCAACCAAGCAAGGAAATGCGGAGCGGGGTGGGAAAGATCAACC 967
QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
Db 968 GATTCAACCAAGTGAACACCTTGCTAAGTTTGCTAATACGGCAATTTATTGCTACAA 1027
QY 281 ArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
Db 1028 CGCAAGCCGAGAGTTGCCCGACGCGCCGCAACCCGTTATTAGATTGATCAAGACAGCG 1087
QY 301 LeuThrProHisProProGlnIleGlnAlaIArgIleValIThrLeuProThrSerValLeu 320
Db 1088 TTGAGCGCCCATCCACCCGCAAAACAGCGGATGTGTGAACATTAACCACTTCACTGTG 1147
QY 321 PheIleAlaGlnYHisAspThrAsnLeuAlaAsnLeuGlnYgIyAlaLeuGlnuLeuAsnTrp 340
Db 1148 TTTATCGCGGACAGATCACTAATCTGGCAATCTCGCGCGGCGCATCTGAGCTCACTCG 1207
QY 341 ThrLeuProGlnYgInProAspAsnThrProProGlnYgIyGluLeuValPheGlnuArgTrp 360
Db 1208 ACGCTTCCCGGTACGCCGATTAACAGCGCGCGAGGTGTGAACGTGTTTGAACGCTGG 1267
QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnIThrLeuGln 380
Db 1268 CGTGGCTTACCGATTAACAGCAAGTGAATCAGGTTTGTGCTGCTTTCAGACTTTACAG 1327
QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlnYgIyValIleLeuThr 400
Db 1328 CAGATGCTGATTAACCCCGCTGCTCATTAATACGCCGCCGAGAGGTGAACGTGACC 1387
QY 401 LeuAlaGlnYsGlnGlnuArgAsnAlaGlnIleMetCysSerLeuAlaGlyPheThrGln 420
Db 1388 CTGGCAGAGATGGAAGACGAAATGCGCAGGCGCATGTGTTGTTGCGAGGTTTACGCAA 1447
QY 421 IleValAsnGlnuAlaArgIleProAlaCysSerLeu 432
Db 1448 ATCGGAATGAAGCAGCATACCGGCGTGCAGTTTG 1483

RESULT 12
US-10-601-319-7
; Sequence 7, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
; TITLE OF INVENTION: AND USING THEM
; FILE REFERENCE: 09010-029011
; CURRENT APPLICATION NUMBER: US/10/601,319
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
```

```
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (188)...(1483)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 403
; OTHER INFORMATION: n = A,T,C or G
US-10-601-319-7

Alignment Scores:
Pred. No.: 5,22e-246 Length: 1901
Score: 2182.00 Matches: 424
Percent Similarity: 98.1% Conservative: 0
Best Local Similarity: 98.1% Mismatches: 8
Query Match: 96.6% Indels: 0
DB: Gaps: 0

US-10-601-319-10 (1-432) x US-10-601-319-7 (1-1901)
QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
Db 188 ATGAAGGAGATCTTAATCCCATTTTATCTCTCGATTCCGTTAAACCCGCAATCTGCA 247
QY 21 PheAlaGlnSerGluProGlnuLeuLeuGlnuSerValIleIleLysSerArgHisGly 40
Db 248 TTCCTCAGAGTGAAGCCGAGAGCTGAAGCTGAAAGTGTGATGTCAATGTCATGCT 307
QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValIThrProAspAlaTrpPro 60
Db 308 GTGCTGCTCCACCAAGACCCAGCACTGATGACGATGTCACCCAGACGATGGCCA 367
QY 61 ThrTrpProValIleLeuGlnYgIyGluLeuThrProArgGlnYgIyGluLeuIleAlaTrpLeu 80
Db 368 ACCTGGCGGTTAAACTGGTTGGCTGACACCGGAGGTGTGAGCTAATCGCTATCTC 427
QY 81 GlyHisTrpTrpArgGlnuArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
Db 428 GGAATTAAACCAACCCAGCGCTGTGTAGCCACGATTTGCTGGGAAAGGGCTGCCCG 487
QY 101 GlnSerGlnYgInValAlaIleIleAlaAspValAspGlnuArgTrpArgLysThrGlnYgIu 120
Db 488 CAGTCTGCTCAAGTCAGCAAGCCAGCACTGATTCGATTCGACAGGCTACCCGTAACAGCGCA 547
QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCCTTCGCGCGCGGCTGGGCTGACCTGATGCAATTAACCGTAACACCCAGGAGATACG 607
QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValIleCysGlnLeuAspAsnAla 160
Db 608 TCCTGCTCCCGATCCGTTAATTAATCTTAATAAATGCGGTTTGCCAACTGAAACCGCG 667
QY 161 AsnValThrAspAlaIleLeuGlnuArgAlaGlyIleSerIleAlaAspPheThrGlyHis 180
Db 668 AACGTGACTGACCGGATCTTACGAGGCGAGAGGCTAATGCTGACTTTACCGGGCAT 727
QY 181 TyrGlnThrAlaPheArgGluLeuGlnuArgValIleuAenPheProGlnSerAsnLeuCy 200
Db 728 CGGCAAAAGCGGTTTCGGAACTGGAACGGGAGTCTTAATTTCCGAATCAAACTGTGCG 787
```

QY 201 LeuLYsArgGLuLYeGLuAspGLuSerCySerLeuThrGlnAlaLeuProSerGluLeu 220
DB 788 CTTAAAGCTGAGAAACAGGACGAAAGCTGTTCAATTAAACGACGCACTTACCACTGGAAATC 847
QY 221 LYeValSerAlaAspCyValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
DB 848 AAGGAGAGCGCGCAATGCTCTCAATTAAACGGTGGGTAAAGCTCGCAATGCTGACG 907
QY 241 GLuLePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyAlaThr 260
DB 908 GAGATATTTCTCTCTCAACACAGGACAGGGAATGCCGAGCCCGGGTGGGAAAGATCAAC 967
QY 261 AepSerHISGlnTyrAsnThrLeuSerLeuHISAsnAlaGlnPheAspLeuGln 280
DB 968 GATTACACCACTGAGAAACCTTGTCTAGTTCATTAACGCCAATTTATTTGCTACAA 1027
QY 281 ArgThrProGluValAlaArgSerAlaThrProLeuLeuAspLeuIleYsThrAla 300
DB 1028 CGCACGCCAGAGTTGCCCGGACGCCGCCCACTTATTAATTGATCAAGACAGG 1087
QY 301 LeuThrProHisProProGlnLYeGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
DB 1088 TTGACGCCCATTCACCGCAAAACAGGCGTATGGTGTGACATTAACCACTTCAGTGTG 1147
QY 321 PheIleAlaGlyHISAspThrAsnLeuAlaLeuGlyGlyValAlaLeuGluLeuAsnThr 340
DB 1148 TTTATCGCGGACAGCATTAATCTGAATCTGGCAATCTCGGCGGCGCATGAGCTCAACTG 1207
QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGlnuArgTyr 360
DB 1208 ACGCTTCCCGGTACGCCGATACACGCCGACAGGTGTGAACTGTGTGAAACGCTGG 1267
QY 361 ArgArgLeuSerAspAsnSerGlnTyrIleGlnValSerLeuValPheGlnThrLeuGln 380
DB 1268 CGTCGCTTACGCAATACAGCAGCATGGAATTCAGTTTCGTCTTCACAGCTTACAG 1327
QY 381 GlnMetArgAspLYeThrProLeuSerLeuAsnThrProProGlyGlyGluValIleLeuThr 400
DB 1328 CAGATGCGTGAATAAACCGCTGTCTTAATTAAGCGCGCCGAGAGGTGAATCACTGACC 1387
QY 401 LeuAlaGlyCySGluGluArgAsnAlaGlnGlyMetCySerLeuAlaGlyPheThrGln 420
DB 1388 CTGGCAGGATGTGAAGACGAAATGCCAGGCGCATGTGTTGCTTGGCAGGTTTTCGGA 1447
QY 421 IleValaGlnGluAlaArgIleProAlaCySerLeu 432
DB 1448 ATCGGATGAAGCACGCATACCGGCGTGCAGTTTG 1483

RESULT 13
US-10-933-115-7
Sequence 7, Application US/10933115
Publication No. US20050281792A1
GENERAL INFORMATION:
APPLICANT: Short, Jay M.
APPLICANT: Kretz, Keith A.
APPLICANT: Gray, Kevin A.
APPLICANT: Barton, Nelson Robert
APPLICANT: Garrett, James B.
APPLICANT: O' Donoghue, Eileen
APPLICANT: Machur, Eric J.
TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: 56462001822
CURRENT FILING DATE: 2004-09-01
PRIORITY FILING DATE: US/09/866,379
PRIORITY FILING DATE: 2001-05-24
PRIORITY FILING DATE: 2000-05-25
PRIORITY FILING DATE: 2000-05-25
PRIORITY FILING DATE: 1999-05-25
PRIORITY APPLICATION NUMBER: US 09/318,528
PRIORITY APPLICATION NUMBER: US 09/291,931

PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 09/259,214
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: US 08/910,798
PRIOR FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 1901
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: CDS
LOCATION: (188)...(1483)
FEATURE:
NAME/KEY: misc_feature
LOCATION: 403
OTHER INFORMATION: n = A,T,C or G
US-10-933-115-7
Alignment Scores:
Pred. No.: 5,226-246 Length: 1901
Score: 2182.00 Matches: 424
Percent Similarity: 98.1% Conservative: 8
Best Local Similarity: 98.1% Mismatches: 8
Query Match: 96.6% Indels: 0
DB: 10 Gaps: 0
US-10-601-319-10 (1-432) x US-10-933-115-7 (1-1901)
QY 1 MetLYsAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
DB 188 ATGAAGGATCTTAATCCATTTTATCTTCTGATTCGGTAAACCCGCAATCTGCA 247
QY 21 PheAlaGlnSerGluProGluLeuLYeGlnValIleValSerArgHISGly 40
DB 248 TTCCTCAGATGTAGCCGGAAGCTGAAAGTGGGATTTGTCACTGATGCTGATGCT 307
QY 41 ValArgAlaProThrLYeAlaIleThrGlnLeuMetGlnAspValThrProAspAlaTyrPro 60
DB 308 GTGCGTGTCCACCAAGGACGACCACTGATGACAGATGTCACCCAGACGATGGCCA 367
QY 61 ThrTyrProValLYeLeuGlyGluLeuThrProArgIlyGlyGluLeuIleAlaTyrLeu 80
DB 368 ACCTGGCCGGTAAACCTGGTGTGCTGACACCGGCGGTGTGAGCTAATGCTTATCTC 427
QY 81 GLYHISYrTTPArgGlnArgLeuValAlaAspGlyLeuLeuProLYeCySGLYCySerPro 100
DB 428 GGAATTTACCAACCGCACGCTGTGTACCGAATGCTGTGGGAAAAAGGCTGCGCG 487
QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLYeThrGlyGlu 120
DB 488 CAGTCTGGTCAAGTCGCGATTAATGCTGATGTCAGACGAGCGTACCCGTAACAGGCGAA 547
QY 121 AlaPheAlaAlaGlyLeuAlaProAspCyValAlaIleThrValHISThrGlnAlaAspThr 140
DB 548 GCCTTGGCCCGCGGCTGGCACCTGACGTGTCAATTAACCGTACATCAACAGAGATACG 607
QY 141 SerSerProAspProLeuPheAsnProLeuYsThrGlyValCYSGlnLeuAspAsnAla 160
DB 608 TCCAGTCCCATCCGTTATTTATCTCTTAATAACCTGGCGTTTGCACATGGAATACCGG 667
QY 161 AsnValThrAspAlaIleLeuGluArgAlaGlyIleSerIleAlaAspPheThrGlyHIS 180
DB 668 AACGTGACTGACGCGATCTCAGACAGGAGGAGGATCAATGCTGACTTAAACGGGCGAT 727
QY 181 TyrGlnThrAlaPheAlaGlyLeuGluArgValAlaLeuAsnPheProGlnSerAsnLeuCYs 200
DB 728 CGGCAAAACGGCGTTTCCGAACTGGAACGGGTGCTTAATTTCCGCAATCAAACTTGTGC 787
QY 201 LeuLYsArgGLuLYeGLuAspGLuSerCySerLeuThrGlnAlaLeuProSerGluLeu 220
DB 788 CTTAAAGCTGAGAAACAGGACGAAAGCTTTCATTAACGACGCACTTACCATCGAAATC 847

```
QY 221 LysValSerAlaAspCysValSerLeuThrGlyValAlaValSerLeuAlaSerMetLeuThr 240
Db 848 AAGGTGAGCGCGCAATGTCATTAACCGGTGCGGAACCTCGCATCATCATCTGACG 907
QY 241 GluIlePheLeuLeuGlnGlnAlaGlnIleMetProGluProGlyTyrGlyValIleThr 260
Db 908 GAGATATTTCTCCCTGCAACAAGACAGGGAATGCGGAGCGGGGTGGGAAGATCACCC 967
QY 261 AspSerIleGlnTyrPheThrThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
Db 968 GATTTCACACGACGTGAACACCTTCTTAAGTTGACATTAACGGCAATTTATTTCTCTCAA 1027
QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleIleValThrAla 300
Db 1028 CGCAGCCGCAAGGTTGCGCGGACCGCGCCGCCCTTATTAAGATTGATCAAGACACG 1087
QY 301 LeuThrProHisProProGlnIleValGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 1088 TTGACGCCCATTCACCGCAAAAACAGGCGTATGTTGACATTACCACTTCAGTGCTG 1147
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 1148 TTTATCGCCGACACGATCTAATCTGCAATCTCGCGCGCGCACCTGAGCTCAACTCG 1207
QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 1208 AGCGTTCGCGTACCGGATTAACAGCGCGGACGAGGTGTGAACGTGTGAACGCTGG 1267
QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1268 CGTGGCGTAACGATTAACAGCGCATGATTCAGGTTTCGCGTCTTCAACACTTTACAG 1327
QY 381 GlnMetArgAspIleThrProLeuSerLeuAsnThrProProGlyGlyValIleLeuThr 400
Db 1328 CAGATGCGTGAATAAACCGCCCTGTCTAATAATACGCGCGGAGAGGTAAACTGACC 1387
QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnIleMetCysSerLeuAlaGlyPheThrGln 420
Db 1388 CTGGCAGGATGTGAAGACCAATGCGGAGGCAATGTTGTTGTGGCAGGTTTACGCA 1447
QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1448 ATCGTGAATGAACGCGCATACCGGCGGTGCA GTTTG 1483

RESULT 14
US-11-056-354-3
; Sequence 3, Application US/11056354
; Publication No. US20050246780A1
; GENERAL INFORMATION:
; APPLICANT: SHORT, Jay M.
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin A.
; APPLICANT: BARTON, Nelson R.
; APPLICANT: GARRETT, James B.
; APPLICANT: O'DONOGHUE, Bileen
; APPLICANT: MAHUR, Eric J.
; TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM AND METHODS
; TITLE OF INVENTION: FOR MAKING AND USING THEM
; FILE REFERENCE: 564462001803/D1370-9C1
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 10/156,660
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
```

```
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli appa phytase
; NAME/KEY: misc_feature
; LOCATION: (403)..(403)
; OTHER INFORMATION: n is any nucleotide
US-11-056-354-3

Alignment Scores:
Pred. No.: 5.22e-246 Length: 1901
Score: 2182.00 Matches: 424
Percent Similarity: 98.1% Conservative: 0
Best Local Similarity: 98.1% Mismatches: 8
Query Match: 96.6% Indels: 0
DB: 14 Gaps: 0

US-10-601-319-10 (1-432) x US-11-056-354-3 (1-1901)

QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 188 ATGAACCGATCTTAAATCCATTTTATCTCTGATTCCTGTAACCCGCAATCTGCA 247
QY 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40
Db 248 TTGCGCTCAGATGAGCGCGAGCTGAACCTGGAAGTGTGATGATTCAGTCCGTCATG 307
QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 411 VALLARGALAPROTHRLYSALATHRGINLEUMETGLNASPVALTMRPROASPAALATRP 60
QY 308 GTGCGTCTCCACCAACCAAGGCGACGCACTAATGCAAGATGTCAACCCAGACGTGGCCA 367
Db 308 GTGCGTCTCCACCAACCAAGGCGACGCACTAATGCAAGATGTGATGATGATGATGATG 367
QY 61 ThrTrpProValLysLeuGlyGluLeuThrProArgIleGlyGluLeuIleAlaTyrLeu 80
Db 368 ACCTGCGCGGTAAACCTGGGTGGCTGACACCGGAGGTGTGAGCTAATCGCTATCTC 427
QY 81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
Db 428 GAACTATTAACCAACCGCGCTGTGTGACGCAAGATGTGCGGAGAAAAGGCGTCCG 487
QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 488 CAGTCTGTCAAGTGTGATTAATGCTGATGTCAAGACGCTACCGCTAATAACGCGCAA 547
QY 121 AlaPheAlaIleGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCCTTCCGCCCGCGGGCTGGCACCTGACTGTGCATAACCGTACATCAACGAGCATACG 607
QY 548 GCCTTCCGCCCGCGGGCTGGCACCTGACTGTGCATAACCGTACATCAACGAGCATACG 607
Db 548 GCCTTCCGCCCGCGGGCTGGCACCTGACTGTGCATAACCGTACATCAACGAGCATACG 607
QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 608 TCCAGTCCCGATCGGTATTAATCCCTTAATAAATCGCGTTGCCAATCGAATAACGCG 667
QY 161 AsnValThrAspAlaIleLeuGluArgAlaGlyCysSerIleIleAlaAspPheThrGlyHis 180
Db 668 AACGTGCTGACGCGGATCTTCAAGGCGAGGAGGAGGTCATATGTGACTTAAACCGG 727
QY 181 TyrGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 728 CGGCAACCGCGTTTCCGGAATGGAACGGGTGCTTAATTTCCGCAATCAAACTTGTC 787
QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 788 CTTAAACGTGAGAAACAGACGAAAGCTGTTCAATTACGACGAGCATTAACATCGGAATC 847
QY 221 LysValSerAlaAspCysValSerLeuThrGlyValAlaSerLeuAlaSerMetLeuThr 240
Db 848 AAGGTGAGCGCGCAATGTCATTAACCGGTGCGGTGAAGCTTCGACATCAATCTGACG 907
```

```

QY 241 GIUIIePheLeuEngInaIaGInGlyMeCProGluPProGlyTTrpGlyArgIleThr 260
| | | | |
Db 908 GAGATATTCTCTCTGCAACAGACAGGAATCCCGAGCCGGGGGGAGAGATCAC 967
QY 261 AapSerHIGlnTrpAanThrLeuSerLeuHISaanaIaGlnPheAapLeuEngIn 280
| | | | |
Db 968 GATTCAACACAGGAGAACACTGTCTAAGTTTCATACCGCAATTTATTTGCTACAA 1027
QY 281 ArgThrProGluValaIaArgSerArgaIaThrProLeuLeuAapLeuIlelyThrAla 300
| | | | |
Db 1028 CGACACCCAGAGGTTCGCCGACGCCGCCGCCCTTTTGAATTTGATCAAGACAGCG 1087
QY 301 LeuThrProHISerProGluInySGInaIaTyrgIValaThrLeuProHISerValleu 320
| | | | |
Db 1088 TTGACGCCCCCATCCACCGCAAAAACAGCGCTATGTGTGACATTACCCACTTCAGTGTG 1147
QY 321 PheIlealIGlyHISaPThrAanLeuAlaAanLeuGlyGlyAlaLeuGluAanTrp 340
| | | | |
Db 1148 TTTATCGCCGACACGATACTATCTGGCAATCTGGGGGGGCACTGGAGGCTCAACTGG 1207
QY 341 ThrLeuProGlyGlnProAapAnThrProProGlyGlyGluLeuValaPheGluatTrp 360
| | | | |
Db 1208 AGCGTTCGGGTCAAGCCGATTAACAGCCGCGAGGTGTGAATCTGGTGTGAAACGCTG 1267
QY 361 ArgArgLeuSerAapAanSerGlnTrpIleGlnValSerLeuValaPheGlnThrLeuGln 380
| | | | |
Db 1268 CGTGGCTAAGCATTAACAGCCAGTGGATTTCAGGTTTCGTGTCTTCAGACTTACAG 1327
QY 381 GlnMeArgAapLySerThrProLeuSerLeuAanThrProProGlyGlyValaIlySerLeuThr 400
| | | | |
Db 1328 CAGATCGGTGATTAACCGCGGTGTCAATTAAATAGCCGCCGAGAGGTGAACCTGAC 1387
QY 401 LeuAlaGlyCySGluGluAArgaanaIaGInGlyMeCySerLeuAlaGlyPheThrGln 420
| | | | |
Db 1388 CTGGAGAGATGTGAAGAGGAATGCGAGGGCATGTGTCTGTGCAAGTTTATCGCAA 1447
QY 421 ILeValaenGluAlaArgIleProAlaCySerLeu 432
| | | | |
Db 1448 ATCGTGATGAAGCAGCATACCGCGGTCAAGTTTG 1483

RESULT 15
US-09-866-379-5
; Sequence 5, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARRON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:

```

```

; NAME/KEY: misc feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-5

Alignment Scores:
Pred. No.: 6,856-246 Length: 1901
Score: 2181.00 Matches: 424
Percent Similarity: 98.1% Conservative: 0
Best Local Similarity: 98.1% Mismatches: 8
Query Match: 96.6% Indels: 0
DB: 3 Gaps: 0

US-10-601-319-10 (1-432) x US-09-866-379-5 (1-1901)

QY 1 MetLyAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
| | | | |
Db 188 ATGAAGAAGCATTTATCCATTTTATCTCTTGTGATTCGTTAAACCCGCAATCTGCA 247
QY 21 PheAlaGlnSerGluProGluLeuLySerValValIleValSerArgHISGly 40
| | | | |
Db 248 TTCGCTCAGAGTGAAGCCGAGCTGAAGCTGGAAGTGTGTGATTTGTCAGTCTATGCT 307
QY 41 ValArgAlaProThrLyAlaIleThrGlnLeuMetGlnaapValThrProAapIaTrpPro 60
| | | | |
Db 308 GTGCGTCTCAACCAAGCCACGCACTGATGCAAGATGTCAACCCAGACGATGGCA 367
QY 61 ThrTrpProValaIlyLeuGlyGluLeuThrProArgIlyGlyGluLeuIleAlaTyrgLeu 80
| | | | |
Db 368 ACCTGCGCGGTAAACTGGGTGACTGAACCGCGNCGTGTGAGCTAATCCCTATCTTC 427
QY 81 GlyHISyTrpArgGlnArgLeuValaIaapGlyLeuLeuProLyCySGlyCySPro 100
| | | | |
Db 428 GGAACATTACCAACGCGCAGGCTGTGTGAGCCGAGATTGCTGCGCAAAAAGGGCTGCCG 487
QY 101 GlnSerGlyGlnValaIleIleIleAlaapValaapGlyuaTrpThrArgLySerGlyGln 120
| | | | |
Db 488 CAGTCTGTCAAGTCCGCAATTAATCTGATGTGACAGCGTAACCCGTAACAGCGCA 547
QY 121 AlaPheAlaIaGlyLeuAlaProAapCyAlaIleThrValHISThrGlnAlaAapThr 140
| | | | |
Db 548 GCCTTCGCGCGGCTGGACCTGACCTGTGCAATACGTACATACCAGGCAATACG 607
QY 141 SerSerProAapProLeuPheAanProLeuLyThrGlyValCySGlnLeuAapAanAla 160
| | | | |
Db 608 TCCAGTCCCGATCCCTTATTTAATCTTAATAAATCGCGTTGCCAATCGATTAACCG 667
QY 161 AenValThrAapAlaIleLeuGluAArgaIaGlyGlySerIleAlaAapPheThrGlyHIS 180
| | | | |
Db 668 AACGTGACTGACGCGATCTCAAGCAGGCGAGAGGTCAATATGTGACTTAAACGGGCAT 727
QY 181 TyrgInThrAlaPheArgGluLeuGluAArgValleuAanPheProGlnSerAanLeuCyS 200
| | | | |
Db 728 CGGCAAGAGGGGTTTCGGGAAGCTGGAACGGGTCTTAATTTCCGCAATCAAACTGTGCG 787
QY 201 LeuLySerArgLyGlyGlnAapGlySerCySerLeuThrGlnAlaLeuProSerGluLeu 220
| | | | |
Db 788 CTTAAACGTGAGAAACAGAGCAAAAGCTGTTCAATTAAACGAGGCACTTACATCGGAATCG 847
QY 221 LySValSerAlaAapCyValSerLeuThrGlyAlaValaSerLeuAlaSerMetLeuThr 240
| | | | |
Db 848 AAGGTGAAGCGCGCAATGTCTCAATTAACCGGTGGGTAAACCTGCAATCAATGTGACG 907
QY 241 GIUIIePheLeuEngInaIaGInGlyMeCProGluPProGlyTTrpGlyArgIleThr 260
| | | | |
Db 908 GAGATATTCTCTCTGCAACAGACAGGAATGCCGAGCCGGGGGGAGAGATCAC 967
QY 261 AapSerHIGlnTrpAanThrLeuSerLeuHISaanaIaGlnPheAapLeuEngIn 280
| | | | |
Db 968 GATTCAACACAGTGAACACCTGTCTAAGTTTCATACCGCAATTTATTTGCTACAA 1027
QY 281 ArgThrProGluValaIaArgSerArgaIaThrProLeuLeuAapLeuIlelyThrAla 300
| | | | |

```

```
Db      1028 CGCAGCGCAGAGGTTGCCGCGAGCGCGCACCCGTTATTGATTGATCAAGCAGCG 1087
Qy      301  LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db      1088 TTGACGCGCCCATCCACCCGCAAAACAGCGGTATGTGTGACATTACCCACTTCAGTGTCTG 1147
Qy      321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db      1148 TTTATCGCCGACACGATTAATCTGGCAAAATCTCGCGGGGCACTGAGCTCAACTGG 1207
Qy      341  ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db      1208 ACGCTTCCCGGTCAAGCCGATTAACAAGCCGCAAGGTGTGAACCTGTGTGAACGCTGG 1267
Qy      361  ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db      1268 CGTCGGCTTAAGCGATTAACAGCAGGTGATTCAGGTTGCTGCTTCAGACTTTACAG 1327
Qy      381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db      1328 CAGATGCGTGAATAAACCGCGCTGTCAATTAAATACGCCGCCCGAGAGGTGAACCTGACC 1387
Qy      401  LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db      1388 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTCGTTGGCAGGTTTACGCAA 1447
Qy      421  IleValaGluGluAlaArgIleProAlaCysSerLeu 432
Db      1448 ATCGTAATGAAGCAGCATACCGGCGTGCAGTTTG 1483
```

Search completed: June 14, 2006, 15:10:36
Job time : 1441.42 secs


```

QY 58 -----AlaTrpPro-----ThrTrpProValIysLeuGly----- 67
DB 313 TCACAGACTGTGGCCAGATCATATGTGTGACTGTCTCTCTTTGGGTCTCCAGAGCGCTTG 372
QY 68 GluLeuThrProArgGlyGlyGluLeuIleAlaTyrLeuGlyYhiIstYrTrpArgGlnArg 87
DB 373 CATCAACACCCCTTAACACAGAAAGTGCAGCCACACTGGGACTCAGAACCAACCAAG 432
QY 88 -----LeuValAlaAspGlyLeuLeu---Pro 95
DB 433 GGAAGAAGACTCAGCGCCCTGGGGTGCCTGGTCTGTGGGACTCAGCATGATGACTTCCAGC 492
QY 96 LysCysGlyCysProGlnSerGlyGlnValAlaIleIleAlaAspValAsp---GlnArg 114
DB 493 TCCTGCGCCTT-CCCCAGCAA-----CTGCTGACTGGGAGCCAGACCGGAGAGCTG 542
QY 115 ThrArgIysThrGlyGlyValAlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrVal 134
DB 543 AGCGACGGGCGCTGGCGAGAGCAAGCTCGGGGTCT----- 575
QY 135 HisThrGlnAlaAspThrSerSerProAspPro-----LeuPheAsnProLeuIysThr 152
DB 576 CACTCAGGACACGAGCCCTCTGCTTGGCCAGGCTTGAGTACTCAGAGCCCTATTCAGGCA 635
QY 153 GlyValCysGlnLeu-----AspAsnAlaAsnValThr----- 163
DB 636 GGAAGTGTCTTCTTGGGATATCGCATTCACCTTAAGATAGAGGAGACTGTGTGACAAAGC 695
QY 164 -----AspAlaIleLeuGlnArg 169
DB 696 TGTGTTCAGCAGCGCTTCCAGAGCCAGAACTAGAGCCAGTGAAGCGACCTGGAGGAG 755
QY 170 AlaGlyIysSerIleAlaAspPheThrGly-----His 180
DB 756 CCTGGA-----TTCTGTGGGTGTCCCGGAGCCACACACAGCCCATGAGC 800
QY 181 TyrGlnThrAlaPhe-----ArgGluLeuGlnArgValIleu 192
DB 801 TACCCAACTGCACTCTTCTCTCTCATCTGTGGCAATGGGGGCCAGGCTTTCGATCTGC 860
QY 193 AsnPheProGlnSerAsnLeuCysLeuIysArgGlyIysGlnAspGlnSerCysSerIleu 212
DB 861 GCCTTCATATGCCAGCGGCTGACACTGTGGCCAAAGTGGCCAGAG----- 905
QY 213 ThrGlnAlaLeuProSerGluLeuIysValSerIleAspCys----- 226
DB 906 ---CAGGTATGAGACACTTATGTTCTGATCTGCTGCTGTGACATCATGTGTCTGAG 962
QY 227 -----ValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThrGlnIlePheLeu 244
DB 963 GAGGTGTAGATCTCTCCGAGCAG---GCCATCCCCCTCCCTTCGAGAA----- 1010
QY 245 LeuGlnGlnAlaGlnGly-----MetProGluProGlyTyrGlyArg 258
DB 1011 CTCAATCGATTGTGATGCTTGGGCTTACAGACCCCTGACAGACGCCCCAGCTGGGGGCG 1070
QY 259 IleThr-----AspSerHisGlnIleTrpAsnThrLeuLeu 269
DB 1071 AGCACCTTACATGAGACGTATGTGTACTTATTCGGTACACAAACACAGGTCTGTAGT 1130
QY 270 Ser---LeuHisAsnAlaGlnPheAspLeuLeuGlnArgThrProGluValAlaArg--- 287
DB 1131 TCTTACGCTGTACACAGATGAGATGACGTCTTGGCCGGAGCCATTGTTGGCCAGTTC 1190
QY 288 -----SerArgAlaThrProLeuLeuAspLeuIleIysThrAlaLeuThrProHis 304
DB 1191 TCTTGGCCAGCAATGTCTTCCAGCCTGTGGTGTGGCCGCGTGCACACCACTCTTAAG 1250
QY 305 ProProGlnIysGln-----AlaTyrGlyValThrLeuProThrSer----- 318
DB 1251 GCCGTAGGAAGAGAGTGAAGCGCCCTTACATGATGTGTTGTGAGAGTCTCCACAGACTGG 1310

```

```

QY 319 -----ValLeuPheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGly--- 333
DB 1311 CAGAGCAAGACGATGCTCTGCTTGGGACTTCAATGTCTATCTGCTTCACTGACCAA 1370
QY 334 -----GlyAlaLeuGluLeu-----AsnTrpThrLeuProGly 344
DB 1371 AAGCGCTTGACAAAGCTGTGAGCTGCGACTGAGCCAGGCTTCCACTGGTGAATGGCCAT 1430
QY 345 GlnProAsnThrProProGlnGlyGly-----LeuVal 356
DB 1431 GGGGAGACACACACATGTCGGGGCCAGACCCCACTGACCTTATGACCGCGTGTGCAC 1490
QY 357 PheGluArgTrpArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPhe 376
DB 1491 GGGGAGGCGCTGCGGAGTCTGTGCACACGCG-----GCTGCTTGAATTC 1538
QY 377 GlnThrLeuGlnIleMetArgAspIysThrProLeuSerLeuAsnThrPro---ProGly 395
DB 1539 CCCACGAGCTTCCAGCTCACCAGGAGAGAGCGCCTCAACATCATGACCACTACCCCGTG 1598
QY 396 GluValIysLeuThrIleAla---GlyCysGlnGluArgAsnAlaGlnGlyMetCysSerI 415
DB 1599 GAGGTGAGCTGAAGACTGAGCCAGGCGCACAGGCTCCAGGCTTCAAGCTTCACTGTCTG 1658
QY 415 euaIaGlyPheThrGlnIleValAsnGluAlaArgIleProAlaCys 430
DB 1659 TTGCTGTATCATCTCTGTCTC-----CCCTCAGCTGTGCCCTGTCTGC 1699

```

RESULT 2

```

US-10-449-902-14217/c
; Sequence 14217, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute for Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449, 902
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 14217
; LENGTH: 1482
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK064835
; DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-14217

```

Alignment Scores:

```

Pred. No.: 0.191 Length: 1482
Score: 110.00 Matches: 66
Percent Similarity: 33.0% Conservative: 33
Best Local Similarity: 22.0% Mismatches: 119
Query Match: 4.9% Indels: 82
DB: 6 Gaps: 11

```

US-10-601-319-10 (1-432) x US-10-449-902-14217 (1-1482)

```

QY 19 SerAlaPheAlaGlnSerGluProGluLeuIysLeuGlnSerValIleValSerArg 38
DB 957 TCGATCTTACCTCTTCCCGCCAGCGTACGCTTCCCTTCTGTGTGTGAGACCGGAAC 898
QY 39 HisGlyValArgAlaProThrIysAlaThrGlnLeuMetGlnAspValThr-ProAspAl 58
DB 897 AGCTGTCTCTGGTGGGCTCCACGACACGCGCTTGTATGTGCGAGATCTCCCTCT----- 843

```

```

Oy 58 aTTPProTtTPProValLlybEuGly-----GluLe 63
Db 842 -TGGGGAACAATGGGAGCAGCCGCTGGATGATGGGCTCGAGACGGAGCAGAGGAGACGAAAC 784
Oy 69 uThrProArgGlyGly-GluLeuIleAlaIleArgLeuGlyIleIstYrTPArgIlnArgLeuV 89
Db 789 TTGGAGTAAAGGAGCCTCCACCTTGTCCACCGCCTCGGACACGAAACGAGTAAACGAGG 722
Oy 89 aLAlaAspArgIleuLeuProLysCyseGlyCyseProGlnSerGlyGln-----104
Db 723 TCGCAGATGCACTG-----GAGTCCCTTCACGGTGGGAGATGCCGTTCACTCC 676
Oy 105 -----ValAlIleIleAla 110
Db 675 AGCGTCCGCTCGGTGGGAGTAAAGGGGCGCGGATGAAGTGAAGCGTGGTCCCTTCTCCG 618
Oy 110 sPValAspGlnArgThrArgLysThrGly-----GluAlaPheAlaAlaGlyLeuAla- 127
Db 615 AGCGTGAACGAGCGGTGAACTCGAGCGGAGCGAGCTGCTGAGCTCTCGAGCGGGTCTCGGCC 555
Oy 128 -----ProAspCybaAlaIleThrValHisThrGlnAlaAspThrSerSerProAsp---144
Db 555 TTCTTGAAGACGTTGTTGTGTAAGCTGCCGCAAGGCGCGCTCGCGGTGAGGACGACG 498
Oy 145 -----ProLeuPheAsnProLeuLysThrGlyValCyseGlnLeuAsp-----AsnAla 161
Db 495 AGGGGACCTCTTGAACGGGGCGGATGCGGATGAAGGGAGGTGATGCTCCGCTCGTGG 438
Oy 161 snValThrAspAlaIleLeuGlnArgAlaGlyLysSerIleAlaAspPheThrGly---179
Db 435 ATCTCGTGGTTCATGTTGTAGACACTCCACAGGCGCTCCGAAAGGGCGGGAGAC 376
Oy 180 -----HisYrGlnThrAlaPheArg 187
Db 375 AGCACCCTCTCTCGCGCCGCGCACCTTGGCCGCGCCACGACCTCATCGCTCCGTG 316
Oy 187 LuLeuGlnArgValLeuAsnPheProGlnSerAlaLeuCyseLysArgGlnLysGln 207
Db 315 ATCTTCTGCTGTGGCGCAC-----GAGCTCG 288
Oy 207 sPGLuSerCyseSerLeuThrGlnAlaLeuProSerGlnLeuLysValSerAlaAspCyav 227
Db 288 ATGCGGCTCGGACGCTCAACGACGAGCACCGCACCACTGAGAGGGCGCGTGTGCGG 229
Oy 227 aLserLeuThrGlyAlaValSerLeuAlaSerMetLeuThrGlnIlePheLeuLeuGln 247
Db 228 CAGCTAGAGGCGGAGCGCGCGCG-----G 205
Oy 247 InAlaGlnGlyMetProGlnLysProGlyTPArgIleThrAspSerHisGlnTrp 265
Db 204 GTGGCCAGGCGGAGGAGACAGGCGGTGAGAGGTGCGAGACACGACATGGCTGG 149

RESULT 3
US-10-449-902-9793/c
Sequence 9793, Application US/10449902
Publication No. US20060123505A1
GENERAL INFORMATION:
APPLICANT: National Institute of Agricultural Sciences.
APPLICANT: Bio-oriented Technology Research Advancement Institution.
APPLICANT: The Institute of Physical and Chemical Research.
TITLE OF INVENTION: Full-Length Plant cDNA AND USES THEREOF
FILE REFERENCE: MOA-A0205Y1-US
CURRENT APPLICATION NUMBER: US/10/449,902
CURRENT FILING DATE: 2003-05-29
PRIOR APPLICATION NUMBER: JP 2002-203269
PRIOR FILING DATE: 2002-05-30
PRIOR APPLICATION NUMBER: JP 2002-383870
PRIOR FILING DATE: 2002-12-11
NUMBER OF SEQ ID NOS: 56791
SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 9793

```

```

; LENGTH: 1519
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK107141
; DATABASE ENTRY DATE: 2002-08-28
US-10-449-9793

Alignment Scores:
Pred. No.: 0.38 Length: 1519
Score: 107.00 Matches: 84
Percent Similarity: 33.4% Conservative: 46
Best Local Similarity: 21.6% Mismatches: 111
Query Match: 4.7% Indels: 148
DB: 6 Gaps: 19

US-10-601-319-10 (1-432) x US-10-449-902-9793 (1-1519)
Oy 15 LeuthrProglInseralaphelialGlnSerGluProgluLeuGlnSerValVal 34
Db 1065 CTGGCTCATCATGTCGCCACGATGCTGCTCAGAGGCC---CTCCCTCTCGCGCGCTGCT 1009
Oy 35 ILeValSerAlaGhIsgIlyValArgAlaProThrIlySalatrhGlnMerGlnAspVal 54
Db 1008 CTGCTCGAAGCCGAGCAGAGTTCCGAGACAGCGCGCGCTCGCGTCTGATCAGACGCT 949
Oy 55 -----ThrAspAla 58
Db 948 GGGATATCTCCATTACGCCCCAGCGGGGCTGAAAGCCACGCTGATCTGCTCCCTCCG 889
Oy 59 TrpProThrTrpProValIyLeuGlyGlnLeuThrProArgIyGlyGlnLeuIleAla 78
Db 888 GCGCGCGCGCTGGCTGCGCGGCGAGCAGAACTGTACCCCGC-----CTCTCT 841
Oy 79 TyrtLeuGly-----HisTyTrrArg----- 85
Db 840 CATCTCGATTGCGCATGTATCACCCGTTGGGCTCTGTGATCTCTCGTGGTCAGACCG 781
Oy 86 -----Gln 86
Db 780 TGGGCTCGCGCGCTTCGCGCGGCTGTTCGGTTGGAAATGAAAGCTCTGATACAGATG 721
Oy 87 ArgLeuValAlaAspGlyLeuLeuProIyCyAsGlyCyProGlnSerGlyGlnValAla 106
Db 720 GAGGAGGTGTGAGC-----CTCCCGTCCGCGGAGGCGCATCTCATGTGCTTTCGTC 667
Oy 107 ILeIleAlaAspVal---AspGluArgThrArgIyThrGlyGlnAlaPheAlaIaGly 125
Db 666 CTCACCTGCCAGGATCTTGACGAAAGGTGAAGAGGAATCTCCCTGCGTGGGCGCC 607
Oy 126 LeuAlaProAspCyValaIleIleThrValHisThrGlnAlaAspThrSerSerProAspPro 145
Db 606 GCCCGCGGTGACTCGGCATA-----GAGCT 580
Oy 146 LeuPheAsnProLeuIyStrhGlyValCyGlnLeuAspAsnAlaAsnValThrAspAla 165
Db 579 CTCGAGAGCAAGACGG---GACTCG 556
Oy 166 IleLeu-----GluArgAlaGlyGly-----SerIle 174
Db 555 GTTCTTCATGAGGAGGAGTCGCGGCGCGGAGACAAGGTGAGCGCCCTTCACAGGCTC 496
Oy 175 AlaAsp-PheThrGlyHisIyTrGlnThrAlaPheAspGlnLeuGlnArgValLeuAsnPh 194
Db 495 GCAAGAGCGCTCGGCGCTCTTGCGGTCCCACTTAAAGAGAACTCGAGATGAAGACGCC 436
Oy 194 eProGlnSerAsnLeuCyAsLeuIyS---ArgGluIyGlnAspGlnSerCySerSerLeu-- 212
Db 435 ATGAGAGAGAGATCATCATGAAGTGTGTGAGAGCGAGCGCGCGCGGCTCTCGCTGA 376
Oy 213 ----ThrGlnAlaLeuProSerGluLeuIyValSerAlaAspCyVal---SerLeuTh 230
Db 375 GCAGGCGCGCGCTCGGCGCTTCAGAGTTTCATCTCTGCACCAAGATGAGCGCGCTGAC 316

```

```
Oy 230 rGlyAlaValSerMetLeuThrGluLeuPheLeuGlnGlnGlnGln 250
Db 315 CGCGCGCA----- 306
Oy 250 yMetProGluProGlyTyrGlyArgIleThrAspSer-----HisGI 264
Db 305 -CTGCCGTCGCGCGCGTGG-----CGGAGACGACGGCCCTGTAGTGTGCACTT 256
Oy 264 nTTPheThrLeuLeu-SerLeuHisAlaGlnPheAspLeuGlnGlnThrProG 284
Db 255 GTGTGCTCCATGTGTCGACGTGGGCGCCGCGCTGTATATAGGGCGCATGACACAC 196
Oy 284 luValAlaArgSerArg----- 289
Db 195 GCGGCGCTGTATCGAGGTCGATTCCTCCGCGACGTTGGCGGAGACCGGAGATGCT 136
Oy 290 -----AlaThrProLeuLeuAspLeuLe-----L 298
Db 135 AATGGCGTCCGCGCGCTCCATTTATCTCTCTAGAGGCGATCCACGTTGCTCCACCA 76
Oy 298 ySThrAlaLeuThrProHisPro 305
Db 75 AGATTCCACGACGCGCATCTCA 53

RESULT 4
US-10-449-902-21879/C
; Sequence 21879, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21879
; LENGTH: 1470
; TYPE: DNA
; ORGANISM: Oryza sativa
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AK072104
; DATABASE ENTRY DATE: 2001-12-06
US-10-449-902-21879

Alignment Scores:
Pred. No.: 0.404 Length: 1470
Score: 106.50 Matches: 70
Percent Similarity: 34.84 Conservative: 15
Best Local Similarity: 23.28 Mismatches: 115
Query Match: 4.7% Indels: 82
DB: 6 Gaps: 13

US-10-601-319-10 (1-432) x US-10-449-902-21879 (1-1470)
Oy 12 LeuIleProLeuThrProGlnSerAlaPheAlaGlnSerGluProGlnLeuLeuGln 31
Db 841 CTTCTCCCGCTCCACGAGTCACTTCCTCTGTGTGTGAGCCGGA-----CAG 791
Oy 32 SerValVal-IleValSerArgHisGlyValArgAlaProThrIysAlaThrGlnLeu 51
Db 790 CTCGTCCTCGGTGGGCTCCACGAC-----ACGCGTTAT 755
Oy 51 tGlnAspValThr-ProAspAlaTrrProThrTrrProValIleuGlnIy----- 67
```

```
Db 754 GTCCAGATCTCCCT-----TGSGCAGATGGGAGGACGCTGTGATGAGGCTC 701
Oy 68 -----GluLeuThrProArgIyGly-GluLeuIleAlaTrrLeuGly 81
Db 700 GGAACGAGCAGAGGAGACGAACTTGAGTGAAGAGAGCTCACTTGTCACCGCTGGA 641
Oy 82 HisTrrTrrArgIlnArgLeuValAlaAspGlyLeuLeuProIyScyGlyCysProGln 101
Db 640 CACGACAGCAGATGAGACGAGGTCCAGATGAGCTG-----GGAGTCTTCAC 593
Oy 102 SerGlyGln----- 104
Db 592 GGTGGGATGCGCTTCACTTCAGAGCTCCGCTCGTGGAGATGAGGCGCGGATGAA 533
Oy 105 -----ValAlaIleIleAlaAspValAspGlnArgThrArgIySerGly-----Glu 120
Db 532 GTAGCGGTGCTGCTTGTGCGAGCTGACGACGCTGACTCGAGCGGACGCTTGAG 473
Oy 121 AlaPheAlaIleGlyLeuAla-----ProAspCysAlaIleThrValHisThrGlnAla 138
Db 472 CTCTCGATGCGGCTCTCGGCTTCTTGAGGACGTTGTGTGAAGCTGCCGAGAGGCC 413
Oy 139 AspThrSerSerProAsp-----ProLeuPheAsnProLeuIySThrGlyValCys 155
Db 412 GCGCTCCGCGTGAAGACGACGAGGCGACCTTCTTGAAGGCGGAGTGGCGGTGAGGCG 353
Oy 156 GluLeuAsp-----AsnAlaAsnValThrAspAlaIleLeuGlnIuArgAlaGlySer 173
Db 352 GAGGTCAATGCTCTCCGTCTGATCTCTGATTCATGTGAGACCTTCACACAGCC 293
Oy 174 IleAlaAspPheThrGlyHisTyrGlnThrAlaPheArg----- 186
Db 292 CTCGAGAAAGGAGGAGGAGACACACCGCTCTGCGCGCGGACCTTGCGCGCGC 233
Oy 187 ---GluLeuGluArgValIleAsnPheProGlnSerAsnLeuCysLeuIySAryGlnIyS 205
Db 232 CACCACTTCATCGCTC-----CGTATCTTGCCTGTGGCGACCA 188
Oy 206 GlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeuIyValSerAlaAsp 225
Db 187 GTCGATCGGCTCCGACGCTCACGACGACGACGACCACTTCGAGGCGCGCTGCT 128
Oy 226 CysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThrGluIlePheLeu 245
Db 127 GCGGACCTAGAGACGAGCGGCGCG----- 101
Oy 246 GlnGlnAlaGlnIyMetProGluProGlyTrrGlyArgIleThrAspSerHisGlnTrr 265
Db 100 ---GRTGGCCAGGCGGAGGAGACGAGCGGTGAGAGGTGCGAGCAGCATGCTGG 44

RESULT 5
US-10-449-902-7723/C
; Sequence 7723, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7723
; LENGTH: 2822
; TYPE: DNA
; ORGANISM: Oryza sativa
```

```

/ PUBLICATTON INFORMATION:
/ DATABASE ACCESSION NUMBER: AKI05071
/ DATABASE ENTRY DATE: 2002-08-28
US-10-449-902-7723

Alignment Scores:
Pred. No.: 1.02 Length: 2822
Score: 106.50 Matches: 70
Percent Similarity: 34.8% Conservative: 35
Best Local Similarity: 23.2% Mismatch: 115
Query Match: 4.7% Indels: 82
DB: 6 Gaps: 13

US-10-601-319-10 (1-432) x US-10-449-902-7723 (1-2822)

QY 12 LeuilePLeuThrProGlnSerAlaPheAlaGlnSerGluProGluLeuLysLeuGlu 31
Db CTTCTCCCGCTCCACGCTCAGCTTCCTCTTGATGTGAGCCGAA-----CAG 2256

QY 32 SerValVal-ILeValSerArgHISgLYValArgAlaProThrIlySalAThrGlnLeu 51
Db CTCTCTCCCGTGGGCGCTCCACGAC-----ACGCCGTGAT 2220

QY 51 tGlnAspValThr-ProPheAlaATrProThrProValLysLeuGly 67
Db GTCCAGATCTCCCT-----TGGGGCATGGGGAGCAGCGCTGTGAATGATGGGCT 2166

QY 68 -----GluLeuThrProArgIlyGly-GluLeuILeAlATyrLeuGly 81
Db GGAGCGGACGAGGAGGAGACGAATCTGAGTAAAGAGAGCTCCACTTGTCCACCGCTCGA 2106

QY 82 HisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProIlybCybGlyCybProGln 101
Db CACCAACAGCGAGTACGAGAGCTGCGAGATGACATG-----GGAGTCTTCAC 2058

QY 102 SerGlyGln----- 104
Db GGTGGGATGCGCTTACCTCCAGCTCCGTCGATGGAGTAAAGGGCGCGGATGAA 1998

QY 105 -----ValAlaIleIleAlaAspValAspGluArgThrArgIlybThrGly-----Glu 120
Db GTAGCGGTGCGCTTCTTGGCGAGCTGACGACGCTGATCTCGAGGCGGAGCTCTGAG 1938

QY 121 AlaPheAlaAgLYLeuAla-----ProAspCybAlaIleThrValHISThrGlnAla 138
Db CTCTCGATGGGGGTCTCGGCTTCTTGAGGACCTGTGTGTGAAGCTGCGCGCAGAGCC 1878

QY 139 AspThrSerSerProAsp-----ProLeuPheAsnProLeuIlybThrGlyValCys 155
Db GCGCTCCCGGTGAGAGCAGACGAGGCGGACCTTTTACGCGGCGGATGCGGGGTAGAGGG 1818

QY 156 GlnLeuAsp-----AsnAlaAsnValThrAspAlaIleLeuGluATGAlaGlySer 173
Db GAGGTGATGTCCTCCGCTGTGAGATCTCTGTTATGTGTAGACACCTCCACAGCGC 1758

QY 174 ILeAlaAspPheThrGlyHISYrGlnThrAlaPheArg----- 186
Db CTCCGAAAGGGGCGGAGAGACACACCGCTCTGCGCGCGCGCGCAGCTTGGCGCGCG 1698

QY 187 ---GluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCybLeuIlybArgGlyLys 205
Db CACCAAGCTTATCGGCTC-----CGATCTTCTGCGTGTGGCAGCGGA 1653

QY 206 GlnAspGluSerCybSerLeuThrGlnAlaLeuProSerGluLeuIlybValSerAlaAsp 225
Db GTGATGCGGCTCGCGACGCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCGCGCTGCT 1593

QY 226 CybValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThrGluIlePheLeuLeu 245
Db GCCGCACTGACGAGCGGAGCGCGCGCG----- 1566

QY 246 GlnGlnAlaGlnGlyMetProGluIlybTrpGlyArgIleThrAspSerHISGlnTrp 265

```

Db	1565	---	GGTGGCCAGGGCGGAGAGACCAAGCGGGGTGGAGAGGTGGGAGACACGACATGTGCTGG	1503
		RESULT 6		
		US-11-189-279-64		
		/ Sequence 64, Application US/11189279		
		/ Publication No. US20060115829A1		
		/ GENERAL INFORMATION:		
		/ APPLICANT: MAO, LI		
		/ APPLICANT: WANG, JIE		
		/ APPLICANT: LIO, WANG		
		/ TITLE OF INVENTION: A METHOD OF TREATING CANCER		
		/ FILE REFERENCE: UTXC:875US		
		/ CURRENT APPLICATION NUMBER: US/11/189, 279		
		/ CURRENT FILING DATE: 2005-07-26		
		/ PRIOR APPLICATION NUMBER: 60/598, 554		
		/ PRIOR FILING DATE: 2004-08-03		
		/ NUMBER OF SEQ ID NOS: 69		
		/ SOFTWARE: PatentIn Ver. 2.1		
		/ SEQ ID NO 64		
		/ LENGTH: 118899		
		/ TYPE: DNA		
		/ ORGANISM: Homo sapiens		
		US-11-189-279-64		
		Alignment Scores:		
		Pred. No.: 211	Length: 118899	
		Score: 106.50	Matches: 84	
		Percent Similarity: 31.1%	Conservative: 33	
		Best local Similarity: 22.3%	Mismatches: 115	
		Query Match: 4.7%	Indels: 144	
		DB: 7	Gaps: 17	
US-10-601-319-10 (1-432) x US-11-189-279-64 (1-118899)				
Qy	61	ThTrPrProAllyLeuGlYluLeuThPrProArgGlyGluLeuLeaIaTyrlau	80	
Db	23020	ACCTGGCTCTGCTTGGCGAAGATTAAAGGGCGAGAGAGTGA--GAGGGTTTCTG	23076	
Qy	81	GLYHsETyTrP-ArgGlnArgLeuValAlaAspGlyLeu-----LeuProly	96	
Db	23077	CCATGCGAGGTGGGGCGTCCGCCCTT---GCACGTGGGATGTGGGGCTTCAGCGCAGGGG	23133	
Qy	96	WYsGlyCySPProGlnSerGlyGlnValaIleIleIleAlaAspValaAspGluArgThr	116	
Db	23134	CTGTCCCTTCCCCCAATCC-----CAGCTTAAAGAAAGAGCTCG	23172	
Qy	116	gLySThngIyglYlualAAlaPheIa---AlaGlyLeuAlaProAspCySAlaIleThrValHi	135	
Db	23173	TGAETGGCTTAAAGCTCGCGGGCGGTGGGGTG-----TGGCCGTGAGAGGGCGCT	23223	
Qy	135	sThngIlnAlaAspThrSerSerProAspProLeuPheAsnProLeuYlThrGlyValCy	155	
Db	23224	GAGCGAAGGGGCGACCTCCCGGGGTGCTCTTAAGCTGAGCCATCAAGCCGGGGAGCAGC	23283	
Qy	155	sGlnLeuAspAsnAlaAsnValThrAspAlaIleLeuGlnArgAlaGlyIserIleAl	175	
Db	23284	GGGGTGGCCCTCTGCGCTGTGCACACCTGGCCCGGAGGAGCGGGCGGGCGCGCG	23343	
Qy	175	aAspPheThngIyHsIeTyrlnThrAlaPheArgGluLeuGluArgValLeuAsnPhePr	195	
Db	23344	TCCCTGTCCCGGCAC-----	23359	
Qy	195	oGlnSerAsnLeuCyLeuLeuYsArgGluYlueGlnAspGluSerCySserLeuThrGlnAl	215	
Db	23360	-----GCCCGCTTACCCTGGGGT	23376	
Qy	215	aLeuProSerGluLeuYsValSerAlaAspCySValSerLeuThrGlyAlaValaSerIe	235	
Db	23377	CTTACCTTCATCTTAAAGCTTAATCCCGGTGTC-----	23410	
Qy	235	uAlaSerMetLeuThngIluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGl	255	
Db	23411	-----GG	23412	

Qy 255 YTPGlyValGileThraSPSerHis----- 263
Db 23413 CTGGCCCACTTCCAGACTGGACACATTAAGCCCTCAAAACCTTAGACACACAGCTCTGG 23472
Qy 264 -----GlnTPraSnThrLeuSerLeuHis----- 272
Db 23473 AAGTAAGCTTCCCGGCTCGTGTGAGAGGGGTGTATGATGCTCCACTTATTCGG 23532
Qy 273 -----AsnIaGlnPheAspLeuGlnIaArgThrProGlnValAlaArgSerAr 289
Db 23533 GGAAGAACTGAAAGCGAAGAAAGACTTAACAAGTTGTAAGACTGCGCTTGGGCCC 23592
Qy 289 GAlaThrProLeuAspLeu----- 296
Db 23593 AGGCTTCCCGCTGAGAGATCTCACTGTGATGAGTTCCAGGCCAGCTTGGCCCGC 23652
Qy 297 -IleuYThrAlaLeuThrProHisProPro-----GlnYsgIaAlaYrgIyVa 313
Db 23653 CCTCCAGCTTTCCTGCTCCCTTTTCTCCCTGTGTAGAGAGAGGCTGCTGGGCTT 23712
Qy 313 1-----ThrLeuProThrSer-----ValLeuPheIleAlaGlyHi 325
Db 23713 CCGTGTTCACCCCTCCCGAAGCTGGAAAGTGTGTAATTAAGGAATTAAGGCGAG 23772
Qy 325 sAspThraSnLeuAlaAsnLeu-----GlyGI 334
Db 23773 AGTATGAACCTTGAAGTCCCTCTCGATTCGATCGCGGTGGTGGCAGACGGGTGG 23832
Qy 334 yAlaLeu-----GlnLeuAsnTrpThrLeuProGlyGlnProAspAsnThrPr 350
Db 23833 CAGTTTGGGAAACAGCTGTATCTCTCTACACAGGCTGCTGGGGCTTCTTGTACCC 23892
Qy 350 0-----ProGlyGlyGlnLeuValPheGlnArgTrpArg 361
Db 23893 CACCTTGAAACGCCAGGTGGCGGACTGATTTTTCAGAGTTGAGA 23938
RESULT 7
US-11-284-978-14
; Sequence 14, Application US/11284978
; Publication No. US20060121508A1
; GENERAL INFORMATION:
; APPLICANT: Garrett-Engle, Philip W.
; APPLICANT: Raymond, Christopher K.
; TITLE OF INVENTION: ALTERNATIVELY SPLICED ISOFORMS OF
; TITLE OF INVENTION: PURINERGIC RECEPTOR P2X, LIGAND-GATED ION CHANNEL, SUBUNITS
; TITLE OF INVENTION: 3 AND 4 (P2RX3, P2RX4)
; FILE REFERENCE: RS0222
; CURRENT APPLICATION NUMBER: US/11/284,978
; CURRENT FILING DATE: 2005-11-22
; PRIOR APPLICATION NUMBER: US 60/633,008
; PRIOR FILING DATE: 2004-12-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 1221
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-284-978-14
Alignment Scores:
Pred. No.: 0.43 Length: 1221
Score: 105.00 Matches: 70
Percent Similarity: 38.1% Conservative: 31
Best Local Similarity: 26.4% Mismatches: 119
Query Match: 4.7% Indels: 46
DB: 7 Gaps: 11
US-10-601-319-10 (1-432) x US-11-284-978-14 (1-1221)
Qy 106 AAlaIleIleAlaAspValaAspGluArgThraGlyThrGlyGlnAlaPheAlaIaGly 125
Db 404 GCTCTGCCGACACCCACAGCAACGAGAGTCTCAACAGGAGGTGCGTACTTCAACGGGT 463

Qy 126 IeuaIaProAspCysAlaIleThrVal-----HisThrgIaAlaAspThrSerSerPro 143
Db 464 CTGTCAAG-ACGTGTAGAGTGGCGGCTGTGCCCCGGTGGAGATGACACACAGCTGCCA 522
Qy 144 AspPro-IeupheAsnProLeu---IystrGlyValCysGlnLeuAspAsnAlaAsnVa 162
Db 523 CAACCTGCTTTTAAAGCTGCGAGAAACTTCACTCTTTTGTGTTAAGAACACATCTGG 582
Qy 162 IThraPAlaIleLeuGlnArgAlaGlyGlySerIleAlaAspPheThrgIyHisYrgI 182
Db 583 TATCCCAATTTAATTTTCAGACAGAGAAATATCTTCCA-----ACATCACACATACT 636
Qy 182 nThr-----AlaPheArgGlnLeuGlnArgValArgValLeuAsnPheProGlnSerAsnLe 199
Db 637 TACTCTAAGTCGTGCTATTTATGATGCTTAATAACAGATCTTCTGCCCATATTCGCTT 696
Qy 199 uCysLeuIyAspGlyIyGlnAsp----- 207
Db 697 GGCAAAATATGAGAAACGAGACACAGATTTCAGACATGGCCGTGAGAGGAGCATC 756
Qy 208 ----GluSerCysSerLeuThrgIaIaLeuProSerGlnLeuIyValSerAlaAspCy 226
Db 757 ATGGGATTCAGGTCAACCTGGAGCTGCAACCTGACAGACCCGCTCTGCT--TG 813
Qy 226 sValSerLeuThrgIyAlaValSerLeuAlaSerMetLeuThrgIuIlePheLeuLeuGI 246
Db 814 CCCAGATATCTTCCTCCGCTCCCTGATATACAGGAGACCTTACGACCAACGATATCTCTGGC 873
Qy 246 ngIaIleGlnIyMetPro-----Glu-ProGlyTPGlyValGileThraSPSerH 263
Db 874 TACAATTTCAAGTTTGGCAAGTACTACAGAGACCTGTGGCAACGAGCGCACGCTC 933
Qy 263 IeGln-----TPraSnThrLeuSerLeu-----H 272
Db 934 ATCAAGGCTATGATGATCCCTTCGACATCATGTGTGGGAAAGCGAGAAATTTGAC 993
Qy 272 IAsnAlaGlnPheAspLeuGlnArgThrProGlnValAlaArgSerArgAlaThrP 292
Db 994 ATCATCCCACTAATGATCAACATCGCTGCTGCGCTGCACTGTAGGCAATGCCCTCTCC 1053
Qy 292 roLeuLeuAspLeuIleIyThraIaLeuThrProHisProGlnIyGlnAlaYrg 312
Db 1054 CA-----CCTTCCCTTCTCAAGACACACCCCTTCAAGTCCAGCTTCT 1098
Qy 312 IyValThrLeuProThrSerValLeuPheIleAlaGlyHisAspThraSnLeuAlaSnL 332
Db 1099 CCCAAGAGATG-----GAGTGCCTTTCATTCGGTAAAGATTCCAGGCTTCTCAGGA 1152
Qy 332 euGIyGIyAla 335
Db 1153 AGGGCACGCA 1163
RESULT 8
US-11-145-307A-189
; Sequence 189, Application US/11145307A
; Publication No. US20060094035A1
; GENERAL INFORMATION:
; APPLICANT: Arcuturus Bioscience, Inc.
; APPLICANT: Erlander, Mark G.
; APPLICANT: Ma, Xiao-Jun
; TITLE OF INVENTION: Identification of Tumors
; FILE REFERENCE: 022041-002020US
; CURRENT APPLICATION NUMBER: US/11/145,307A
; CURRENT FILING DATE: 2005-06-03
; PRIOR APPLICATION NUMBER: US 60/577,084
; PRIOR FILING DATE: 2004-06-04
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 189
; LENGTH: 2711
; TYPE: DNA
; ORGANISM: Homo sapiens

US-11-145-307A-189

Alignment Scores:

Pred. No.:	1,34	length:	271
Score:	105.00	Matches:	87
Percent Similarity:	34.9%	Conservative:	51
Best Local Similarity:	22.0%	Mismatches:	154
Query Match:	4.7%	Indels:	104
DB:	7	Gaps:	16

US-10-601-319-10 (1-432) x US-11-145-307A-189 (1-2711)

Qy	26	ProGluLeuValbLeuGlnSerValValIleValSerArgHisGlyValArgAlaProThr	45
Db	239	CCGGAAGTGTCCCGCAGCGCATCTGTAGACTCGGCCACAGAGTTTAAGCCCTCGAC	298
Qy	46	LysAlaThrGlnLeu-----MetGlnAspValThrProAspAlaTrp	59
Db	299	ATCTCTCCCAAGCTCCGGTCACGCATGCTCGTCCAGCAAGATCCTTGACAGTACTAC	358
Qy	60	ProThrTrpProValLysLeuGlyLeuLeu-----ThrPro	71
Db	359	GAGACTGGCAGACATCCGGCTGGAGTGAATGAAGGGGGCTCCAGGCCAAGGTGGCACCC	418
Qy	72	ArgGlyGlyLeuLeuAlaTyrLeuGlnGlyHisTyrTrpArgGln-----	86
Db	419	AAG-----GTGTGGAGAAAGATTGGGAGACTACAAACGCCAAGAACCTTACATGTTT	469
Qy	87	-----ArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro	100
Db	470	GCTTGGAGATCCGAGACCGGCTCTGGGTGAGGGGTGTGTGACAAATACACTGTGGCC	529
Qy	101	GlnSerGlyGlnValAlaIleIleAlaAspValaAspGlnArgThrArgIysThrGlyGln	120
Db	530	AGTGTCACTCCATTAATAGATCATTC-----CGGAC---AAAGTGCAGCA	574
Qy	121	AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrVal-----	134
Db	575	CCATT-----AACCTCCCTATGACACAGCTGGCCACCAAGTCCCTGATCCCGGA	628
Qy	135	HisThrGlnAlaAspThrSerSerProAspProLeuPheAsnProLeuIysThrGlyVal	154
Db	629	CACACGGCGATCCCGACGCTCACTGTTAACTCCCGGAGTCAACC-----	673
Qy	155	CysGlnLeuAspAsnAlaAsnValThrAspAlaIle-Leu---GluArgAlaGlyGlySe	173
Db	674	---CAGTGGAGATTCCCTGGGCTCCACCTACATCAATAGGGCTCCTGGCATTCGTGAG	730
Qy	173	IleAlaAspPheThrGlyHisTyrGlnThrAlaPheArg-----	186
Db	731	CTGTGCAGCCGACAAAGAGAAATGATGACAGTGTACAGATAGCTGCCGACTAAGCATT	790
Qy	187	-----GluLeuGlnArgValLeuAsnPheProGlnSerAs	198
Db	791	GACTCACAGACAGACAGCGGACCCGAAAGCACTTTCGACGGAGTCTTCACCGAG	850
Qy	198	nLeuCysLeuLysArgGlyLysGlnAspGlnSerCysSerLeuThrGln-----	214
Db	851	CACCACTCGAGCGCGTCAAGTGGCCATTGTGAGGGGACGACATACCAGAGGCTATGCG	910
Qy	215	-----AlaLeuProSerGlyLeuLysValSerAlaAspCysValSerLeuThrGlyAl	232
Db	911	TCCCCAGCCACACCCAAAGCGAGAGGAGGCTCTACCCGCTGCCCTTGTCAACAGACC	970
Qy	232	aValSerLeuAlaSerMetLeuThrGlnIlePheLeuLeuGlnGlnAlaGlnGlyMetPr	252
Db	971	CTGAGACGACGGGA-----AGGCCAACCCTGACCCCTTTC	1007
Qy	252	oGluProGlyTyrGlyArgIleThrAspSerHisGlnTrpAsnThrLeuSerLeuHis	272
Db	1004	AAACAGCCACTGGGGCGCAACTCTCG--ACTTACCAAGACCTACCCCGTGTGGCAGATCC	1066
Qy	272	AsnAlaGlnPheAspLeuLeuGlnArgThrProGluValAlaArgSerArgAlaThrPr	292

Db	1063	TCACGACCCCTGGCCATCAATGAAGGAAACCCCGAGGATGTCGAGTTCTAGCTCCACCCC	1122
Qy	292	o-----leuLeuArgleuIleuYThrAlaLeuThrProHisr	305
Db	1123	TTGCTCTTATCTAGCTCCGCGCTTTGAGATG-----	1156
Qy	305	oProGlnlyvgAlaIaTyrGlyValThrLeuProThrservalLeuPhe-----	321
Db	1157	-----CAGCAAGTCGGCTCCGGGGTCCGCGCTTCGAATGCTTTCCCATGCTGC	1206
Qy	322	-----IleAlaGlnIleAspThrValnleuAlaValnleuGlyAlaValnleuGluLeuValnThr	340
Db	1207	CTCCGGTACGCGGCAATTCACGGGCGCAGGCGCTCTCTCAGCGGCGAGAAATGATGGGGCC	1266
Qy	340	rThrIleuProGlnIleProAspValnThrGyrProGluIleGlu	354
Db	1267	CAGCTGTCCGAGTACCCACCCGACCTCCGACCGAGGACAG	1309

RESULT 9
MS-10-50

```

US-10-501-834-1/c
; Sequence 1, Application US/10501834
; Publication No. US20060088828A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Peter C., Ward, Christopher J., Rossetti, Sandro, and Torres,
; APPLICANT: Vincente B.
; TITLE OF INVENTION: Polycystic Kidney Disease Nucleic Acids
; TITLE OF INVENTION: and Proteins
; FILE REFERENCE: 07039/386US1
; CURRENT APPLICATION NUMBER: US/10/501,834
; CURRENT FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: PCT/US03/02038
; PRIOR FILING DATE: 2003-01-23
; PRIOR APPLICATION NUMBER: 60/351,110
; PRIOR FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 221
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 12225
; TYPE: DNA
; ORGANISM: Homo sapien
; US-10-501-834-1

```

Alignment Scores:	
Pred. No.:	12.8
Score:	104.50
Percent Similarity:	34.9%
Best Local Similarity:	23.3%
Query Match:	4.6%
DB:	6
Length:	12235
Matches:	109
Conservative:	54
Mismatch:	167
Indels:	139
Gaps:	24

US-10-601-319-10 (1-432) X US-10-501-834-1 (1-12225)

QY	1	MethylalalLeuileuileProheleuSerleuileuileProleuthrProglSerla	20
		:::	
Db	4439	ATGAAAAGAGTGAATTCCCTGACACTGCTGTGATGCCATGACACAGGACGTGACG	4380
QY	21	PhealGlnSerIupProgluLeuLysleuGlnSerValValIleValSerArgHisGly	40
		:::	
Db	4379	TTTCAGGAGAGAAAGAGTCCA-----GGC	4355
QY	41	ValArgAlaProThrLysValaThrGlnLeuMetGlnaSerValThrProaPalaTrpPro	60
		:::	
Db	4355	AAAGGCTCACCTCCCAAGCTAAC-----TGGCAG	4322
QY	61	-----ThrTrpProValLysLysleuGlnLeuThrProArgLysLysGlu-----	75
Db	4325	AGAAATGATGTGCTCTCCCAAACTCAAAATGACACAAAGTAAGAAAGACCCGAGAGTCAAC	4266
QY	76	-----LeuIlealaTrpLeuGlnHisIeTrpArgGlnArgLeuValAlaAspGly	92
		:::	
Db	4265	CGAAGTCACTCTCTTCAAGATA-----AGAAAGCAACCCCTCACTAGTAAGT	4218

```
QY      93  LeuLeuProLysCysGlyCysProGlnSerGlyGlnValAlaIleIleAlaAspValAsp 112
Db      4217 ATGCTCCCAACCAKATGCCGAACCTGCGCATGGAGATGGCCATTTATCCAGGCACTCAT 4158
QY      113  GluArgThrArgLysThrGlyGlnAlaPheAlaIleGlyLeuAlaProAspCysAlaIle 132
Db      4157 GCAATATGCTGAGACACCAACAGCATATATAGCA-----AATCCATCTGCTT-CTG 4108
QY      133  ThrValHisThrGlnAlaAspThrSerSerProAspProLeuPheAsnProLeuLysThr 152
Db      4107 ACGTACTTGGAGAGATAGATGCCAGC-----CTCCGCACT 4072
QY      153  -----GlyValCys----- 155
Db      4071 GTGAAGAGGATGAGATCCAGACAGGCTACGTCCTCGAAGAGCTGTGCTCAC 4012
QY      156  -----GlnLeuAspAsn-----AlaAsnValThrAspAlaIleLeu 167
Db      4011 ATCACAGTTCAAGTTCCCAAGAGAGATACAGATTGGAGAGATTACTTCTCCCAACATG 3952
QY      168  GluArgAlaGlyGlySerIleAlaAspPheThrGlyHis-----TyrGlnThrAla 184
Db      3951 CAG---GCTCAGCGCTGCTATTGTGATTCTCTTCCTTCATGGCAGTCACTACTGCTGTC 3895
QY      185  PheArgGluLeuGlu-ArgValIleuAsnPheProGlnSerAsnLeuCysLeuLysArgG1 204
Db      3894 TGCCGCTTCATACATGAAAGTGAAAGCTTTCC-----AC 3859
QY      204  uLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGluLeuLysValSerAl 224
Db      3858 CAAGCTTGTGAGAGACCAAGCAGGCGAGAGAACCTGTTCACCCAGCCTCCAGCGCAGC 3799
QY      224  aAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThrGluIlePheLe 244
Db      3798 TGGAACAGTG----- 3789
QY      244  uLeuGlnGlnAlaGlnGlyMetProGluProGly---TyrGlyArgIleThrAspSerHi 263
Db      3788 -----GGAGCGCCCGCATCGGGATATCTGGGGGGCTGGCAGAGTTTCA 3745
QY      263  sGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuGln---ArgTh 282
Db      3744 CCAG-----ATGCTCGCTCCGTTAAGTTCAACATGTCACAGACCGCAT 3700
QY      282  rProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuLysThrAlaLeuTh 302
Db      3699 GCCCAACAATACCAACCAAGCTGGGCTCCCTGTAAGCCTATTCTGTGATGCTAG 3640
QY      302  rProHisProProGlnLysGln---AlaTyrGlyValThrLeuProThrSerValLeuPh 321
Db      3639 GATGGTCCCTCCAGCAGAGGACCAACAGAGGCTCCATGCTGAAAACTTCTGTGAGGTA 3580
QY      321  eIleAlaGlyHisAsp-----ThrAsnLeuAlaAsnLeuGlyG1 334
Db      3579 CTGATGTGAGATCAACCCCTGTGAGTGAATGCTGACCCCATGATAGAGAC-GGAAA 3521
QY      334  yAlaLeuGlnLeuLeuThrThrLeuProGlyGlnProAspAsnThrProProGlyGlyG1 354
Db      3520 TTCTGTGAGACCAAGCTGGCAGTGG--GGGAGAGCCACCTCCAGGCCCA---AGCG 3467
QY      354  uLeuValPheGlnArgTrpArgLeuSerAspAsnSerGlnTrpIleGlnValSerLe 374
Db      3466 ACTGTGTGTGAACCGGAGACCAAGGATCTGTGACTGTGACATTCACATCCAAATCCGAT 3407
QY      374  uValPheGlnThrLeuGlnGlnMetArgAspLys--ThrProLeuSerLeuAsnThrPro 394
Db      3406 AGTTCATCAG-----CTGCGCACTCAATGACCAAGGCTTCACCGC 3365
QY      394  roGly-----GluValLysLeuThrLeuAlaGlyCysGlnGluLysArg--- 407
Db      3364 CTGCTATATGCTTATATGTTTCTGCTCAGATGCAATATACTGATTTAAGAAAGAGCAT 3305
QY      408  -----AsnAlaGlnGly 411
```

```
Db      3304 ATGTAATGCTCTGGGA 3288
RESULT 10
US-10-511-937-358
; Sequence 358, Application US/10511937
; Publication No. US20060088836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, Macdonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 50661200104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 358
; LENGTH: 3646
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-511-937-358

Alignment Scores:
Pred. No.: 4.37 Length: 3646
Score: 101.50 Matches: 124
Percent Similarity: 32.9% Conservative: 59
Best Local Similarity: 22.3% Mismatches: 195
Query Match: 4.5% Indels: 181
Gaps: 29

US-10-601-319-10 (1-432) x US-10-511-937-358 (1-3646)
QY      2  LysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAlaPhe 21
Db      412  CGAGTCTCTTTTGGACAGGTAGGCTCGCGGCTCCCACTATCATCAAGCCATG 471
QY      22  AlaGlnSerGluPro---GluLeuLysLeu-----GluSerValIleValSer 37
Db      472  GGTGGAGCCAGCCAGGGGAACTTCAGATCAGCTGGGAGAGAGCCAGCTCCAGAAATCAGT 531
QY      38  -----ArgHisGlyValArgAlaProThrLysAlaThr----- 48
Db      532  GATTTCGAGGTACGAATCTCGCTATGCCCCAGAGATCCAAAGAACTCACTGCTCC 591
QY      49  -----GlnLeuMetGln----- 52
Db      592  ACGGTATACAGCTGATTTGGCCAGAAACCTGCTGCCCTGTGCAGAGCCCTCACTCA 651
QY      53  -----AspValThrProAspAlaTrpProThrTrpProValLysLeuGly--- 67
Db      652  GCCTTCCTGTGACCAAGTCTCCATGCTTCAAGCCCAATAGCCCTGGCAATGAGCA 711
QY      68  -----GluLeuThrProArgGlyGlyGlu---Leu 76
Db      712  AAGCAGACTCCCAAGATAGAAAGCTTCAAGCTGACAGAGAGGGTGGAACTGCTC 771
QY      77  IleAlaTyrLeu-----GlyHisTyrTrpArgGln---ArgLeuValAlaAspGly 92
Db      772  ATTCAGAGACTCCAGCTGGCAACTCTACTGCTGAGCTGGCAGCTGGCAGCAAGCACTGATGG 831
```

```

QY 93 LeuLeuProLyseCyseGlyCyseProGlnSerGlyGlnValAlaIleIlealaSerValAsp 112
DB 832 ---ATCTCCCTCGGTGCTCTGGGAGATCTGGTCCCTCCCTGTGACCTGACCTGG--- 885
QY 113 GlnArgThrArgIleThrGlyGlnValAlaIleGlyLeuAlaProAspCyseAlaIle 132
DB 886 ---CCTGGAGATGACAGGCGACCTTGGACCTG--- 912
QY 133 ThrValIleThrGlnAlaSerThrSerSerProAspProLeuPheAsnProLeuLysThr 152
DB 913 ---CAATGCTTTACC----- 924
QY 153 GlyValCyseGlnLeuAspAsnAlaAsnValThrAspAlaIleLeuGlnArgAlaGly 172
DB 925 ---TTGACCTGGAAGATGTTATCTGTCATGTGACGACGACGACCATGCT 972
QY 173 SerIleAlaAspPheThrGlyIleGlyGlnThrAlaPheArgGlyLeuGlnArgValLeu 192
DB 973 AGCTCCCAAGGCTTCTTCTTACCAACAGAGGCGACGCGTGTGCCCCAGACAGAG--- 1026
QY 193 AsnPheProGlnSerAsnLeuCyseLeuLysArgGlnLys----- 205
DB 1027 ---TACCCATCTGGGAGAACTCGAAGAGAGAAAGAAACAATCCAGACTACAGACC 1083
QY 206 ---GlnAspGlnSerCyseSerLeuThrGlnAlaLeuProSerGlyLeuLysValSerAla 224
DB 1084 CCACAGTTCTCTCGGCGACCTTCAAGTACAGAAAGACAGCATTTATTCATCTCTGG 1143
QY 225 AspCyseValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThrGlnIlePheLeu 244
DB 1144 GAGGTGACCAAGCCCCGGGTACTGTTAC-----AGCTACCTGGGCTCCCCCTTCTGG 1197
QY 245 LeuGlnGlnAlaGlnGlyMetProGlnProGly-----TrrGlyArgIleThrAspSer 262
DB 1198 ATCCACAGAGGCTGTGGGCTCTCCCAACCCAACTTGCACTGAGGAGAGTCTCCAGTGG 1257
QY 263 His-----GlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeu 279
DB 1258 CATCTGGAATTGGAGTGGAGAC-----CCATGCTCTGGGCGAGCCCAAA-GACTGTGA 1313
QY 280 GlnArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThr 299
DB 1314 TCA---ACTCCG---ATACACAGAGAGAGGCGCATGAGACTGAGAGGCTGAGAGCGCG 1367
QY 300 AlaLeuThrProHisPheProGlnLysGlnAlaIleGlyValThrLeuProThrSerVal 319
DB 1368 TCTCGGGCGCCGAGAGAGGACCTTGAGCTGCGCGCGCATCTCGTACCG----- 1418
QY 320 LeuPheIleAlaGly-----HisAspThrAsnLeuAlaAsnLeuGly 333
DB 1419 ---TTTACAGCTGGCGCGCAAGGCTCAAGGCGCCCACT-ACCAAGGCTCTGGAGCTGCT 1474
QY 334 GlyAlaLeuGlnLeuAsnTrpThrLeuProGlyGlnProAspAsnThrPro----- 350
DB 1475 GGTGCGACCAACTGAGTGGAGACCGGCAACGAGCGCGCGATCTCTTGGTGAACG 1534
QY 351 -----ProGlyGlyLeuValPheGlnArgTrp--- 360
DB 1535 CTGTGATCTAGTGGCGCTCAAGCGGCTGCTG-GGCTGCTGCTGCTGAGTGGGAG 1593
QY 361 -----ArgArgLeu----- 363
DB 1594 TTTCTGCACTACAGAGACTGAGGATGCGCTGTGGCCCTCACTTCAGACCTGCAC 1653
QY 363 ----- 363
DB 1654 CGGCTCTAGGCGCACTTCTAGGACACTGCAAGCGCTGAGCGCGCCCAAGGCGCACATGC 1713
QY 364 SerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGlnMetArg 383
DB 1714 TCAAGTACTCTGTGAGAGAGACCCAGGCTCTT---GAATCTCTCCCAAGTCTCA 1770
QY 384 AspLysThrProLeuSerLeuAsnThrProProGlyGlnVal----- 397

```

```

DB 1771 GAGAGGACTCTTGGCCCTGTGTCTCTCCAGGCGCCAAATGACTACAGGATTGCAG 1830
QY 398 -----LysLeuThrLeuAlaGlyCyseGlnArgAsnAlaGlnGlyMet 412
DB 1831 CCTTCTGCGCGGAGCATGCCCCCTGTGTGTGGCCCAACCATGCTGATGACAGGCTCC 1890
QY 413 CyseSerLeuAlaGlyPheThrGlnIleValAsnGlnAlaArgIlePro 428
DB 1891 TGTGTACC-----ACCACATTTGCCAACATTCCTACTTACCA 1929

RESULT 11
US-11-293-697-315
; Sequence 315, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: Patencin Ver. 2.1
; SEQ ID NO 315
; LENGTH: 2408
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-293-697-315

Alignment Scores:
Pred. No.: 4.17 Length: 2408
Score: 99.00 Matches: 65
Percent Similarity: 30.5% Conservative: 21
Best Local Similarity: 23.0% Mismatches: 90
Query Match: 4.4% Indels: 106
DB: 7 Gaps: 13

US-10-601-319-10 (1-432) x US-11-293-697-315 (1-2408)
QY 5 LeuIleProPheLeuSerLeuLeuIle-----ProLeuThrProGlnSerAla 20
DB 255 CTGCTTCCAGAGCTCCGACCCCAAAATTCCTCTCAAGACGACCGGCGGAGCGCTGC 314
QY 21 PheAlaGlnSerGlnProGlnLeuLysLeuGlnSerValIleValSerArgHisGly 40
DB 315 AAAGCTGTCTCCGCGCTCACCTG-----GCTGCGCGC 347
QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
DB 348 CGGCGGTCCCTTGAAAGCGCGCGAG-----CAGATCGGGGTCTCCGCGCTGGGGGC 401
QY 61 ThrTrpValLysLeuGlyGlnLeuThrProArgGly-----GlyGlnLeuIleAla 78
DB 402 GTTGGCTTTTGTTCCTCCGCGCGAGGTCCGGCCAGAGTGGCGTTGGAGCTCCGCGCGC 461
QY 79 TyrLeuGly-----HisTyr 83
DB 462 GGCCTCGGAGCAAAAGCGCGGAGCGGAGCGTGCAGAGACTTCAACCAAGCCCGCTCGC 521
QY 84 TrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCyseGlyCyseProGlnSerGly 103
DB 522 ACCGAAAAAGCAAGACCGCGGAGCGGAGCGGAGCGGAGCGCTGC----- 569
QY 104 GlnValAlaIleIleAlaAspValAspGlnArgThrArgLysThrGlyGlnAlaPheAla 123
DB 570 -----TGGCTGAGAGAACGACGCGCAACGAGCGCAAGCGGCG- -GCGAAAGG 617
QY 124 AlaGlyLeuAlaProAspCyseAlaIleThrValHisThrGlnAlaAspThrSerSerPro 143
DB 618 GCAGGCGGAGAGACC-----ACCCTGAGCGCGGCTCTCCAGCGGCGCTCTCT 665

```



```

; GENERAL INFORMATION:
; APPLICANT: Inohara, Naohiro
; TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
; FILE REFERENCE: UM-06967
; CURRENT APPLICATION NUMBER: US/11/317,329
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US/10/128,174
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 2696
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-11-317-329-2

Alignment Scores:
Pred. No.:      8.44      Length:      2696
Score:          96.50     Matches:      104
Percent Similarity: 33.1%   Conservative: 53
Best Local Similarity: 21.9%   Mismatches: 167
Query Match:      4.3%     Indels:      152
DB:              7        Gaps:       22

US-10-601-319-10 (1-432) x US-11-317-329-2 (1-2696)

QY      8 PheLeuSerLeuLeuLeuProLeuThrProGlnSerAlaPheAlaGlnSerGlu----- 25
DB      1207 TTTCTGACATTTACATTCAGACAGACATATCTGCTGCTGACAGTGTGCTGTGGCA 1266
QY      26 ----ProGlnLeuLeuLeuSerValIleValSerArgHisGlyValArgAlaPro 44
DB      1267 GTCCCAAGACAGACAGGCGCTGCGAGAAAGTGTCTCTGCAGAGCTGTGCGCCAGCC 1326
QY      45 ThrLyAla-----ThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
DB      1332 AAGGAGCTTAATGAGACATCAACCCAGAACTGATGACAGTCACTCAGAAACTACCTG 1386
QY      61 ThrTrpProValLeuLeuGlyGlu-----LeuThrProArgGlyGlyGlnLeuIle 77
DB      1387 AAGCGGGCGCTTCAGCTCCGACCGTAAGAAATTGTGCTCCGAGATGAGGAATCTGTG 1446
QY      78 AlaTrpLeuGlyHisTrpTrp---ArgGlnArgLeuValAlaAspGlyLeuLeuProLys 96
DB      1447 ATCTAATGAGAACAAAGGTCAACCCCTCCCAATTCCTGTGTGCGCCAGGCGAGTGGAG-- 1503
QY      97 CyGlyCyProGlnSerGlyGlnValAlaIleIle-----AlaAspValAsp 112
DB      1504 -----CAGGTGAGGTGTGCTGCGCCACGAGGTAGACGTGGAC 1542
QY      113 GlnArgThrArgLysThrGlyGlnAlaPheAlaAlaGlyLeuAlaPro----- 128
DB      1543 TGCAGACG-----GCCCTCGAATACACGCCCTCCCTCGATCGCC 1581
QY      129 -----AspCyAlaAlaIleThrValHisThrGlnAlaAspThrSer 141
DB      1582 GCCCAGACCAACCAACCGACCTGTGTGCTGCTTGTGCAATGTGCTGATGCAAC 1641
QY      142 SerProAsp-----ProLeuPheAsnProLeuLysThrGlyValCyssGln 156
DB      1642 CGAGTGGATGAGAGATGCTGGCGGCCCACTTGACGCCCAAGATGGAGATGACCCC 1701
QY      157 -----LeuAspAsnAlaAsnValThrAspAlaIleLeuGlnArgAlaGly 171
DB      1702 ACTGCGCGCTGCTGCTGACCAACGGGCGCTGTGTGGATGCC---CAGGAACGTGAAGGG 1758
QY      172 GlySerIleAlaAspPheThrGlyHisTrpGlnThrAlaPheArgGlnLeuGlnArgVal 191
DB      1759 TGGACCCCTCTTCACCTGCTGCA-----CAGAAATACCTTGAAGATGTGGACGCGCTT 1812
QY      192 LeuAsnPheProGlnSerAsnLeuCyLeuLysArgGlnLysGlnAspGlnSerCyssSer 211

```

```

DB      1813 CTGTGTCCCGTCAGGCTGACCCCAACCTGGCTGAG----- 1848
QY      212 LeuThrGlnAlaLeuProSerGlnLeuLysValSerAlaAspCyValSerLeuThrGly 231
DB      1849 -----GCTAAGGGCAAGAACCCCTTCATGTGGCGCTTACTTTGGC 1890
QY      232 AlaValSerLeuAlaSerMetLeuThrGlnIlePheLeuLeuGlnIleGlnIleMet 251
DB      1891 CATGTTAGCTGTGTCAGCTGCTGAC----- 1917
QY      252 ProGlnProGlyTrpGlyArgIleThrAspSerHisGlnTrpAsnThrLeuSerLeu 271
DB      1918 -----AGCCAGGGGGCTGAGTTGGATGTCTCAGCAGAGAAACCTGGAACACACACG 1968
QY      272 HisAsnAlaGlnPheAspLeuLeuGlnArgThrProGlnValAlaArgSerArgAlaThr 291
DB      1969 CACCTGGCA-----GTAGAGCGGGGCAAAAGTGAAG 1998
QY      292 ProLeuLeuAspLeuIleLeuThrAlaLeuThrProHisAspProGlnLysGlnAlaTrp 311
DB      1999 GCCATTCAAACACCTGTGAAGTGAAGCGGTCCCTGATGCCCTTGACCAAGAGCGTAT 2058
QY      312 Gly----- 312
DB      2059 GGCCTCAGTGCACATCTGACGCTGCCAGGGGCAATACCTGATCTGCAAGATGCTGCTCAGG 2118
QY      313 -----ValThrLeuProThr-----SerValLeuPheIleAlaGly 324
DB      2119 TACGAGACCAAGCCTTGAAGCTGCCACCCACAGGCTGGACACCCCTGATCTAGACAGCC 2178
QY      325 HisAspThrAsnLeu-----AlaAsnLeuGlyGly 334
DB      2179 TACAAAGGCGCACTGAGATCATCTCATCTGTCGACAGAGAGCCAGCAAAATGGGTGCT 2238
QY      335 AlaLeuGlnLeuAsnTrpThrLeuProGlyGlnProAspAsnThrProProGlyGlyGln 354
DB      2239 CTGGAGCGTGAATCTGAGATCTCC-CTTGCA---CTTAGTGAACCCACGGGAGAGA--- 2291
QY      355 LeuValPheGlnArgTrpArgArgLeuSerAsnSerGlnTrpIleGlnValSerLeu 374
DB      2292 -----GGCGGTGGTGTGACGACTCTGCA 2315
QY      375 ValPhe-----GlnThrLeuGlnGlnMetArgAspLysThrProLeuSerLeuAsn 391
DB      2316 GTGTGGGGCTGACCCCAATGTCTGACAGACGACGCTGACACACCT----- 2363
QY      392 ThrProProGlyGlnValLysLeuThrLeuAlaGlyCyssGlnGlu 406
DB      2364 -----CACTGTGGCGGTCCAGAGAGACCTTCTGAGTGTCTATCAA 2405

RESULT 14
US-11-317-329-4
; Sequence 4, Application US/11317329
; Publication No. US20060105413A1
; GENERAL INFORMATION:
; APPLICANT: Inohara, Naohiro
; TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
; FILE REFERENCE: UM-06967
; CURRENT APPLICATION NUMBER: US/11/317,329
; CURRENT FILING DATE: 2005-12-22
; PRIOR APPLICATION NUMBER: US/10/128,174
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2696
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-11-317-329-4

Alignment Scores:
Pred. No.:      8.44      Length:      2696

```

Score:	96.50	Matches:	104
Percent Similarity:	33.1%	Conservative:	53
Best Local Similarity:	21.9%	Mismatches:	167
Query Match:	4.3%	Indels:	22
DB:	7	Gaps:	15

US-10-601-319-10 (1-432) x US-11-317-329-4 (1-2696)

[illegible][illegible]

RESULT 15

Sequence 5, Application US/11317329

Publication No. US20
; GENERAL INFORMATION:

APPLICANT: Nunez, G

1. TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling

CURRENT APPLICATION NUMBER: US/11/317,317

PRIOR APPLICATION NUMBER: US/10/128.174

PRIOR FILING DATE: 2002-04-23

; SOFTWARE: PatentIn version 3.1

LENGTH: 2696

TYPE: DNA
ORGANISM: Homo sapiens

US-11-317-329-5

Alignment Scores:

FILED NO.:
Score:

Percent Similarity
Best Local Similarity

Query Match:

US-10-601-313-10

8 Phe
111
QY

Db 1207 TTT

26 --- QY

Db 1267 GTC

45 7th Ave

2
3
4
5
6
7
8

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

```

QY      61  ThrTrpProValIleuEnglyGu-----LeuThrProArgGlyGlyIleuIle 77
Db      1387  AAGCGGGCCCTTCAGCTCCGACCGTAAAGATTTCGCCGAGAAATGAGAACTGTGT 1446
QY      78  AlaTYrLeuGlyHisTYrTrp---ArgGlnArgLeuValAlaAspGlyLeuLeuProLys 96
Db      1447  ATCTATGAAACAAGGTCAACCCCTCCAAATTCCTGTGTGGCCAGGGCACTGTGGAG--- 1503
QY      97  CySGlyCyAspProGlnSerGlyGlnValAlaIleIle-----AlaAspValAsp 112
Db      1504  -----CAGGTGAGGTTCCTGTGCGGCCACAGAGAGTACGTGAGAC 1542
QY      113  GluArgThrArgIleThrGlyGlyAlaPheAlaAlaGlyLeuAlaPro----- 128
Db      1543  TGCACAGACG-----GCCCTGTGATACACGCGCCCTCTCATGACGC 1581
QY      129  -----AspCyValAlaIleThrValHisThrGlnAlaAspThrSer 141
Db      1582  GCCCAGGACCAAGCAACCCGACCTCTGTGCTCTTGTGGACATGTGTCTGATGCCAAC 1641
QY      142  SerProAsp-----ProLeuPheAsnProLeuIleThrGlyValCySGln 156
Db      1642  CGAGTGGATGAGAGATGCTGGGCGCCACCTTGACAGCCAGAAATGGGATGACCGC 1701
QY      157  -----LeuAspAsnAlaAsnValThrAspAlaAlaIleuGlyIleGly 171
Db      1702  ACTGGCGGCTGTCTCTGAGCAACGCGGCGCTGTGTGATGCC---CAGGAACTGTAAAGG 1758
QY      172  GlySerIleAlaAspPheThrGlyHisTYrGlnThrAlaPheArgGlyIleuGlyArgVal 191
Db      1759  TGGACCCCTCTTCACCTGCTGCA-----CAGAAATACCTTGAGAAATGTGGACGGCTT 1812
QY      192  LeuAsnPheProGlnSerAsnLeuCyLeuIleValArgGlyIleAspGlyIleAspGlySer 211
Db      1813  CTGGTCTCCCGTCAAGGTGACCCCAACCTGCGTAG----- 1848
QY      212  LeuThrGlnAlaLeuProSerGlyLeuIleValSerAlaAspCyValSerLeuThrGly 231
Db      1849  -----GCTAAGGGCAAGACCCCTCATGTGGCGCGCTTATTTGGC 1890
QY      232  AlaValSerIleuAlaSerMetLeuThrGlyIlePheLeuEnglyGlnAlaGlnIleMet 251
Db      1891  CATGTTAGCTGTGTCAAGCTGCTGAC----- 1917
QY      252  ProGlnProGlyTYrGlyArgIleThrAspSerHisGlnTrpAsnThrLeuLeuSerLeu 271
Db      1918  -----AGCCAGGGGGCTGAGTTGATGCTCAGCAGAAACCTGAGAAACACCACTG 1968
QY      272  HisAsnAlaGlnPheAspLeuLeuGlnArgThrProGlyValAlaArgSerArgAlaThr 291
Db      1969  CACCTGGCA-----GTAGAGCGGGGCAAAAGTGAAG 1998
QY      292  ProLeuLeuAspLeuIleIleThrAlaLeuThrProHisProProGlnIleGlnAlaTYr 311
Db      1999  GCCATCCAAACACTGCTGAAGGTGAGCGGTCCCTGATGCCCTTGACCAAGCGGCTAT 2058
QY      312  Gly----- 312
Db      2059  GGGCCCACTGCACTGCAAGTGCAGGGGCAAAATACCTGATCTGCAAGATGCTGCTCAGG 2118
QY      313  -----ValThrLeuProThr-----SerValLeuPheIleAlaGly 324
Db      2119  TACGAGACCAAGCTTGAGCTGCCACCCACAGGGCTGAGCACCCCTGCATCTAGCAGCC 2178
QY      325  HisAspThrAsnLeu-----AlaAsnLeuGlyGly 334
Db      2179  TACAAAGGCGCACTGAGATCATCATCTGCTGGCAGAGAGCCAGCAACATGGGTGCT 2238
QY      335  AlaLeuGlyLeuAsnTrpThrLeuProGlyGlnProAspAsnThrProProGlyGlyGly 354
Db      2239  CTTGAGAGCTGTGAATGAGACTCC-CCTGCA---CCTAGCTGACCGCACGGGGAGGA--- 2291

```

```

QY      355  LeuValPheGluArgTrpArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeu 374
Db      2292  -----GGCGGTGGTGTCAAGCACTGCTGCA 2315
QY      375  ValPhe-----GlnThrIleuGlnIleMetArgAspIleThrProLeuSerLeuAsn 391
Db      2316  GTGTGGGCTGACCCCAATGCTGCAGAGCATGCAAGCTGTGACACCCCT----- 2363
QY      392  ThrProProGlyGlyValIleLeuThrIleuAlaGlyCySGlyGly 406
Db      2364  ---CCACCTGGCGGTCCAGAGAGACCTTCTGTAGTGTATCA 2405

```

Search completed: June 14, 2006, 14:25:58
 Job time : 98.6927 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 14, 2006, 11:46:02 ; Search time 4895.67 Seconds
(without alignments)
7401.587 Million cell updates/sec

Title: US-10-601-319-10

Perfect score: 2258

Sequence: 1 MKALIPFLSLIPPLPQSA.....CSLAGPTQVNEARIPACSL 432

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

MODEL=frame+pan.model -DEV=xlp
-Q=/abss/ABSSWEB_spool/US10601319/runat_12062006_180052_27516/app_query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=rbt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=ads803p
-OUTTMP=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=ads803p
-USER=US10601319 -CGEN_1_1_12067 @runat_12062006_180052_27516 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCL=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est3:*
3: gb_est4:*
4: gb_est5:*
5: gb_est6:*
6: gb_hic:*
7: gb_hic2:*
8: gb_est7:*
9: gb_est8:*
10: gb_est9:*
11: gb_ges1:*
12: gb_ges2:*
13: gb_ges3:*
14: gb_ges4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1016	45.0	853	13	CL662734 PRI0142b
2	1001	44.3	612	3	BU074127
3	905	40.1	529	3	BU713770
4	888	39.3	746	4	CA093060
					SCCCL200

c	5	847	37.5	595	5	CF326092
c	6	839	37.2	895	8	CN762997
c	7	778	34.5	868	8	CN754382
c	8	768	34.0	864	8	CN757565
c	9	623	27.6	383	7	AM036132
c	10	623	27.6	383	7	AM036134
c	11	621	27.5	393	7	BE520240
c	12	583.5	25.8	531	3	BU713769
c	13	522	23.1	706	8	CN759004
c	14	463	21.4	354	1	AA545747
c	15	443	19.6	707	2	BU619443
c	16	389.5	17.2	1175	9	CX943192
c	17	368	16.3	320	7	BE364043
c	18	365.5	16.2	821	13	CL655831
c	19	360.5	16.0	842	13	CL65745
c	20	335	14.8	208	2	BM409016
c	21	335	14.8	208	2	BM412806
c	22	328	14.5	546	2	BU040796
c	23	298	13.2	757	4	CA093061
c	24	288	12.8	657	2	BG457132
c	25	284	12.6	1171	8	CV672454
c	26	267	11.8	179	7	BE461872
c	27	258	11.4	508	2	BU030203
c	28	239.5	10.6	480	1	AL586973
c	29	229	10.1	772	14	DX010757
c	30	227	10.1	470	9	DN201839
c	31	225	10.0	254	7	BE520241
c	32	219.5	9.7	692	5	CK564106
c	33	192.5	8.5	551	5	CK535965
c	34	185.5	8.2	687	10	DV079873
c	35	185	8.2	732	9	CX700688
c	36	183.5	8.1	734	9	CK507686
c	37	182.5	8.1	575	5	CK543282
c	38	177.5	7.9	696	1	AU004475
c	39	154	6.8	769	8	CO386885
c	40	153	6.8	863	8	CV672453
c	41	147	6.5	1044	6	CNS080M3
c	42	144.5	6.4	645	2	BU037937
c	43	140	6.2	250	14	DX065699
c	44	132.5	5.9	2201	6	CR926080
c	45	132	5.8	1430	6	CR615614

ALIGNMENTS

RESULT 1	CL662734	853 bp	DNA	linear	GSS 09-JUL-2004
LOCUS	PRI0142b_B11	(853)	Mixed stage fosmid library of P. pacificus var. California	Pristionchus pacificus genomic, genomic survey sequence.	
DEFINITION	CL662734				
ACCESSION	CL662734				
VERSION	CL662734.1				GI:50150877
KEYWORDS	GSS				
SOURCE	Pristionchus pacificus				
ORGANISM	Pristionchus pacificus				
REFERENCE	Bukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Neodiplogasteridae; Pristionchus.				
AUTHORS	Srinivasan J., Otto G.W., Kahlow U., Geisler R. and Sommer R.J.				
TITLE	(bases 1 to 853)				
JOURNAL	Apapadri: an Acedb database for the nematode satellite organism Pristionchus pacificus				
COMMENT	Nucleic Acids Res. 32 (1), D421-D422 (2004)				
	14681447				
	Contact: Sommer RJ				
	Evolutionary Biology				
	Max-Planck-Institute for Developmental Biology				
	Spemannstr. 37-39, Tuebingen D-72076, Germany				
	Tel: 00497071601371				
	Fax: 00497071601498				
	Email: ralf.sommer@uebingen.mpg.de				
	This library was generated at Caltech, Pasadena, USA and end				
	sequenced at Vancouver, Canada.				

Seq primer: T7
 Class: fosmid ends.
 Location/Qualifiers
 source
 1. 853
 /organism="Pristionchus pacificus"
 /mol_type="Genomic DNA"
 /strain="California"
 /db_xref="taxon:54126"
 /clone_lib="Mixed stage fosmid library of P. pacificus
 var. California"
 /note="Vector: pGpifos-5 Fosmid vector"

ORIGIN

Alignment Scores:
 Pred. No.: 3,43e-92 Length: 853
 Score: 1016.00 Matches: 193
 Percent Similarity: 98.5% Conservative: 0
 Best Local Similarity: 98.5% Mismatches: 3
 Query Match: 45.0% Indels: 0
 DB: 13 Gaps: 0

US-10-601-319-10 (1-432) x CL662734 (1-853)

Oy 237 SerMetLeuThrGlnIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyr 256
 Db 2 TCATGCTGACGGAGATATTTCTCTCGCAACAGCAAGGAATCCGGAGCCGGGGTGG 61
 Oy 257 G1yArg1IeThrAspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnDhe 276
 Db 62 GGAAGAGTACCGGATTCACACAGTGAACACCTTGCTAAGTTTCATTAACGCGCAATT 121
 Oy 277 AspLeuLeuGlnArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeu 296
 Db 122 TATTTGCTACACGACGACGAGGTGTCGCGACGCCGCCATCCGCTTATGTGATTG 181
 Oy 297 IleYsThrAlaLeuThrProHisProProGlnYsGlnAlaYsGlyValThrLeuPro 316
 Db 182 ATCATGGACGGCTTACGCCCATTCACCGCAAAACAGCGCTATGTGTCATTAACC 241
 Oy 317 ThrSerValLeuPheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeu 336
 Db 242 ACTTCAGTACTGTTTATTTGCGGACACGATCTACTATCTGGCAATCTCGCGCGGCATCTG 301
 Oy 337 GluLeuAsnTrpThrLeuProGlyGlnProAspAsnTrpProProGlyGlyGlnLeuVal 356
 Db 302 GAGCTCACTGACCTTCCGCTCCGCTACGCGATTAACGCGCGAGTGTGAACCTGATG 361
 Oy 357 PheGluArgTrpArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPhe 376
 Db 362 TTTGACCTGCGCGCTGCTAAGCATTAACAGCCAGTGGATTCAAGTTTCGCTGATCTTC 421
 Oy 377 GlnThrLeuGlnGlnMetArgAspYsThrProLeuSerLeuAsnThrProProGlyGln 396
 Db 422 CAGACTTACACGACATCGGTGATTAACGCGCGCTGATTAATACCGCGCGCGAGAG 481
 Oy 397 ValYsLeuThrLeuAlaGlyCYsGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 416
 Db 482 GTGAACAGACCTTGGCAGAGTGAAGAGCGCAAAATGGCGCAGGCGATGTGTTGTCGA 541
 Oy 417 GlyPheThrGlnIleValAsnGlnAlaArgIleProAlaCysSerLeu 432
 Db 542 GGTTTACGCAATCGTGAATGAACGACCATACCGGCGTGCAGTTTG 589

RESULT 2
 BU074127/c 616 bp mRNA linear EST 29-SEP-2003
 LOCUS BU074127 NIBB Mochii normalized Xenopus tailbud library Xenopus
 DEFINITION laevis cDNA clone XL090106 5', mRNA sequence.
 ACCESSION BU074127
 VERSION BU074127.1 GI:17504316
 KEYWORDS EST.
 SOURCE Xenopus laevis (African clawed frog)
 ORGANISM Xenopus laevis

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 Xenopodinae; Xenopus; Xenopus.
 1 (bases 1 to 616)
 Kityama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-i, T. and
 Kohara, Y.
 Expressed genes in X. laevis embryo
 Unpublished (2001)
 Contact: Tadasu Shin-i
 Center for Genetic Resource Information
 National Institute of Genetics
 111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: cchini@genes.nig.ac.jp
 The information of this clone is available through the following
 URL.
 http://xenopus.nibb.ac.jp.
 FEATURES
 source
 1. 616
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="XL090106"
 /tissue_type="whole embryo"
 /dev_stage="stage 25"
 /clone_lib="NIBB Mochii normalized Xenopus tailbud
 library"

ORIGIN

Alignment Scores:
 Pred. No.: 6,95e-91 Length: 616
 Score: 1001.00 Matches: 198
 Percent Similarity: 96.6% Conservative: 0
 Best Local Similarity: 96.6% Mismatches: 7
 Query Match: 44.3% Indels: 0
 DB: 2 Gaps: 0

US-10-601-319-10 (1-432) x BU074127 (1-616)

Oy 89 ValAlaAspGlyLeuLeuProLYsCysGlyCysProGlnSerGlyGlnValAlaIle 108
 Db 615 GTACCCGACGATGCTGCGCAAAAGGGCTGCCGCGAGTCTGATCGAGTCCGATATT 556
 Oy 109 AlaAspValaAspGluArgThrArgYsThrGlyGlnAlaPheAlaIleGlyLeuAlaPro 128
 Db 555 GCTGATGTGACGAGCGCTACCCGTTAAACAGCGCAACCTTCGCGCGGCTGGCACT 496
 Oy 129 AspCysAlaIleThrValHisThrGlnAlaAspThrSerSerProAspProLeuPheAsn 148
 Db 495 GACTGTGCATTAACCGTACATACCCAGGCAAGTACGTCCAGTCCGATCCGTTATTAT 436
 Oy 149 ProLeuYsThrGlyValCYsGlnLeuAspAsnAlaAsnValThrAspAlaIleLeuGln 168
 Db 435 CCTCTAAAACTGGCGTTTCCGATCACTGATTAACCGCAACGATCAACGCACTCCTCAGC 376
 Oy 169 ArgAlaGlyIleSerIleAlaAspPheThrGlyHisYsTrpGlnThrAlaPheArgGlnLeu 188
 Db 375 AGGCGAGAGGGTCAATTTGCTGACTTAAACCGGCATCGGCAAAAGCGGTTCGGAACTG 316
 Oy 189 GlnArgValLeuAsnPheProGlnSerAsnLeuCysLeuYsArgGlyLeuGlnAspGln 208
 Db 315 GAAGGGTGTCTTAATTTTCCGCAATCAACTTGTGCTTAAACGTGAGAAACAGACGA 256
 Oy 209 SerCysSerLeuThrGlnAlaLeuProSerGlnLeuYsValSerAlaAspCysValSer 228
 Db 255 AGCTGTTCAATTAACGACGAGCATTAACATCGGAATCAAGGTGAGCGCGCAATATGTCGA 196
 Oy 229 LeuThrGlyAlaValSerLeuAlaSerMetLeuThrGlnIlePheLeuLeuGlnAla 248
 Db 195 TTAACCGGTGCGTGAAGCTTCGATCAATGCTGACGAGAGATATTTCTTCGCAACAAGCA 136
 Oy 249 GlnGlyMetProGluProGlyTyrGlyArgIleThrAspSerHisGlnTrpAsnThrLeu 268

```

Db      125 CTGGGAATCGGAGCGGGTGGGGAAGATCAACGATTCAACACAGGAAACCTTG 76
Qy      269 LeuSerLeuH1AaSnA1aGlnPheAaPLeuLglnA1aThrProGluValA1aArgSer 288
Db      75 CTAAATTTGGCAATAGACCGCAATTTATTTGTTCAACAGCACCGCAGAGGTGCCCCGAC 16
Qy      289 ArgAlaThrProLeu 293
Db      15 CCGCCGACCCCGTTA 1

RESULT 3
LOCUS   BU713770
DEFINITION  BU713770 529 bp mRNA linear EST 23-OCT-2003
          SUTABUC02 Adult SJC 7/94 Schistosoma japonicum cDNA similar to
          pdb1IDRP1A Chain A, Crystal Structure Of Phytase Complex Of
          Escherichia Coli Phytase At Ph 6.6. Phytase Is Bound With Its
          3-Phosphate In The Active Site. Hg2+ Cation Acts As An
          Inter-molecular Bridge, mRNA sequence.
ACCESSION  BU713770
VERSION   BU713770.1 GI:28321126
KEYWORDS  EST.
SOURCE    Schistosoma japonicum
ORGANISM  Schistosoma japonicum
          Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
          Strigeidae; Schistosomatidae; Schistosomatidae; Schistosoma.
          1 (bases 1 to 529)
          Hu, W., Yan, Q., Shen, D.K., Liu, F., Zhu, Z.D., Song, H.D., Xu, X.R.,
          Wang, Z.J., Rong, Y.P., Zeng, L.C., Wu, J., Zhang, X., Wang, J.J.,
          Xu, X.N., Wang, S.Y., Fu, G., Zhang, X.L., Wang, Z.Q., Birdley, P.J.,
          McManus, D.P., Xue, C.L., Feng, Z., Chen, Z., and Han, Z.G.
          Evolutionary and biomedical implications of a Schistosoma japonicum
          complementary DNA resource
          Nat. Genet. 35 (2), 139-147 (2003)
          12973349
JOURNAL   Contact: Zeguang Han
          Chinese National Human Genome Center at Shanghai
          351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
          201203, P. R. China
          Tel: 86-21-50801919 (ex. 45)
          Fax: 86-21-50801922
          Email: hanzg@chgc.sh.cn.
COMMENT   Location/Qualifiers
          1..529
             /organism="Schistosoma japonicum"
             /mol_type="mRNA"
             /strain="Chinese (Anhui) strain"
             /db_xref="taxon:6182"
             /sex="Male and female"
             /tissue_type="whole body"
             /dev_stage="Adult worms"
             /lab_host="Mouse and rabbit"
             /clone_lib="Adult SJC 7/94"
             /note="Vector: lambda ZAP-II XR; Site 1: EcoR I; Site 2:
             XhoI I; Several hundred adult Schistosoma japonicum
             (Anhui, P.R. China, strain), of mixed sex, were perfused
             from the mesenteries of experimentally infected mice and
             rabbits at the Queensland Institute of Medical Research,
             Brisbane, Australia (QIMR), and stored for several months
             in liquid nitrogen. Subsequently, mRNA was isolated at the
             QIMR from lysates of these worms by oligo dT
             chromatography, using a kit from Pharmacia. The mRNA was
             then shipped to Clontech, Palo Alto, CA, USA, who
             constructed a cDNA library. First strand synthesis was
             primed with an oligo-dT-XhoI-primer and synthesized using
             M-MuLV reverse transcriptase. Second strand synthesis was
             accomplished with RNase H and T4 DNA polymerase. The
             double stranded cDNA was ligated to EcoRI linkers,
             digested with EcoRI and XhoI, and ligated into the
             phagemid vector lambda ZAP II XR. After construction of
             this directional library by Clontech, it was returned to
             the QIMR. During analysis of the library at the QIMR, we
             have found that a small percentage, 2% to 3%, of the

```

clones contain inserts that appear to be highly homologous to sequences from salmonid fishes, as determined by homology comparisons using BLAST and by Southern hybridization analysis to genomic DNA from salmon (Sigma Chemical Co., St. Louis, MO) under stringent washing conditions. The remainder of the clones appear to contain S. japonicum sequences."

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Score:	Matches:	Conservative:
3,126-81	529	905.00	174	0
Percent Similarity:		98.9%		0
Best Local Similarity:		98.9%		2
Query Match:		40.1%		0
DB:		3	Gaps:	0

US-10-601-319-10 (1-432) x BU713770 (1-529)

```

Qy      224 A1AaPcyValSerLeuThrG1yAlaValSerLeuA1aSerMetLeuThrGluLeuPhe 243
Db      1 GCCGACAAATGCTCAATTAACCGTGGGTAAAGCTTCGACATCAATGCTGACGAGATATT 60
Qy      244 LeuLeuGlnGlnA1aGlnGlyMeCProGluPProGlyTProGlyAArgI1eThraPserHis 263
Db      61 CTCCTGCAACAAGCAACAGGAAATGCCGAGCCGGGTGGGAAGATCACGATTCACAC 120
Qy      264 GlnTPAaThrLeuLeuSerLeuH1aSnA1aGlnPheAaPLeuLglnA1aArgThrPro 283
Db      121 CAGTGAACACCTTGCTAAAGTTTGATTAACGGCAATTTTATTTGCTCAACAGCACGCCA 180
Qy      284 GluValA1aArgSerA1aThrProLeuLglnA1aThrProGluValA1aThrPro 303
Db      161 GAGGTGGCCCGAGCGCGGCCACCCCGTTATTGAATTCACAGACAGCGTTGACGCC 240
Qy      304 HisProProGlnLysGlnA1aTyrg1yValThrLeuProThrSerValLeuPhe1leA1a 323
Db      241 CATCCACCGCAAAAACAGCGTATGCTGATGACATTAACCATTCAGTCTGTTTATCGCC 300
Qy      324 GlyHisAaPThrAaSnLeuA1aSnLeuGlyGlyAlaLeuGlnLeuAaSnTrpThrLeuPro 343
Db      301 GGAACAGATACATTAATCTGGCAAAATCTCGCGCGGACAGAGCTCACTGAGCGCTTCC 360
Qy      344 GlyLInProAaPThrProProGlyGlyGlnLeuValPheGlnA1aArgTrpArgArgLeu 363
Db      361 GGTGACCGGATTAACACCGCCGACAGTGTGAACGTGTTGAACGCTGGCGTGGCTA 420
Qy      364 SerAaPaAaSerGlnTrpI1eGlnValSerLeuValPheGlnTrpLeuGlnGlnMetArg 383
Db      421 AGCGATTAACAGCAAGTGAATTCAGGTTTGGCTGCTTCCAGACTTTACAGCAGATGGCT 480
Qy      384 AspLysThrProLeuSerLeuAaSnThrProProGlyGlnValValLeu 399
Db      481 GATTAAGCGCGCTGTCTTAAATACGCGCGCCGAGAGGTGAACCTG 528

RESULT 4
LOCUS   CA093060 746 bp mRNA linear EST 23-SEP-2003
DEFINITION  SCCCL2001D10.b CL2 Saccharum officinarum cDNA clone SCCCL2001D10
          3', mRNA sequence.
ACCESSION  CA093060
VERSION   CA093060.1 GI:34946367
KEYWORDS  EST.
SOURCE    Saccharum officinarum
ORGANISM  Saccharum officinarum
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
          clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
          complex.
          1 (bases 1 to 746)
          da Silva, A.L., da Silva, F.R., Kemper, E.L., and Arruda, P.
          The libraries that made SUCSEST
          JOURNAL  Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

```


QY 310 AATYTGlyValThrLeuProThrSerValLeuPheIleAlaGlyHisAspThrAsnLeu 329
DB 475 GGTATGATGTGACATTACCACTTCAAGTGTCTTTATCCCGGACACACTACTATCTG 416
QY 330 AAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrpThrLeuProGlyGlnProAspAsnTrp 349
DB 415 GCMAATCTCGCGCGCCACTGGAGCTCAACTGACCTTCCCGGTCAAGCCGATACACG 356
QY 350 ProProGlyGlyGluLeuValPheGluArgTrpArgLeuSerAspAsnSerGlnTrp 369
DB 355 CCGCCAGGATGTGACTGATGTTTGAACCTGCGCTGAAGCATTAACGCCAGTGG 296
QY 370 IlegInValSerLeuValPheGlnThrLeuGlnGlnMetArgAspIleThrProLeuSer 389
DB 295 ATTCAGGTTTCGTGTCTTCCAGACTTTACAGCAGATGCGTATTAACGCCGCTGTCA 236
QY 390 LeuAsnTrpProProGlyGlyGluValIleValLeuThrLeuAlaGlyCysGlnGluArgAsnAla 409
DB 235 TTAAATACCGCCCGGAGAGGTAAGTGAACCTTGCAGATGTGAAGAGGGAATGCG 176
QY 410 GlnGlyMetCysSerLeuAlaGlyPheThrGlnIleValAsnGluAlaArgIleProAla 429
DB 175 CAGGCGATGTGTCTGTTGCGCAGGTTTACGCAATCTGTAATGAAGCAGCAGTACCGGCG 116
QY 430 CysSerLeu 432
DB 115 TGCAGTTTG 107

RESULT 6
LOCUS CN762997/1 895 bp mRNA linear EST 20-MAY-2004
DEFINITION ID0AA5DC09RM1 ApMs Acyrthosiphon pisum cDNA clone ID0AA5DC09 5',
ACCESSION CN762997.1 GI:47536920
VERSION CN762997.1 GI:47536920
KEYWORDS EST.
SOURCE Acyrthosiphon pisum (pea aphid)
ORGANISM Acyrthosiphon pisum
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.
1 (bases 1 to 895)
Hunter, W., Martinez-Torres, D., Rahbe, Y., Sabater-Munoz, B.,
Stern, D., Tagu, D. and Wincker, P.
An expressed sequence tags database for the pea aphid Acyrthosiphon
pisum
Unpublished (2004)
JOURNAL
COMMENT Contact: D. Tagu
INRA Rennes
UMR BIO3P BP 35327, F-35653 Le Rheu Cedex France
Tel: +33.2.23.48.51.65
Fax: +33.2.23.48.51.50
Risk of contamination by bacterial sequences from obligatory
(Buchnera) or facultative endosymbionts. These sequences were
obtained in the frame of the International Consortium of Aphid
Genomics in collaboration with Genoscope
PCR primers
FORWARD: CAGGAAACAGCTATGACC
Plate: 5 row: C column: 9.
Location/Qualifiers
1..895
/organism="Acyrthosiphon pisum"
/mol_type="mRNA"
/cultivar="developmentstage"
/db_xref="taxon:7029"
/clone="ID0AA5DC09"
/tissue_type="whole insect"
/dev_stage="nymphs and adults (parthenogenetic females)"
/lab_host="XLA Blue"
/clone_id="ApMs"
/note="Vector: pBS-SK minus, site 1: EcoRI, site 2: XhoI;
Sample name: ID0AAA ; Plant growth place: Department of

Ecology & Evolutionary Biology, Princeton University ;
Soil conditions: Soil ; Sowing date: 01/06/1999 ; no stress ;
Harvesting date: 01/06/1999 ; Stress date: 01/06/1999 ;
Description: Aphids inoculated on one-week old Vicia faba
under non-sterile conditions. All parthenogenetic stages
and both winged and wingless adults were collected for
library construction. ; experimental condition: long
photoperiod (16-hr light/8-hr dark at 18 c)"

ORIGIN
Alignment Scores:
Pred. No.: 3,576-74 Length: 895
Score: 839.00 Matches: 160
Percent Similarity: 99.4% Conservative: 0
Best Local Similarity: 99.4% Mismatch: 1
Query Match: 37.2% Indels: 0
DB: 8 Gaps: 0

US-10-601-319-10 (1-432) x CN762997 (1-895)

QY 272 HIsAenAlaGlnPheAspLeuGluAlaArgTrpProGluValAlaArgSerArgAlaThr 291
DB 894 CATTAACGGCAATTTATTTGCTTACACGACGCGAGGTTGCCCGCAGCGCCGAC 835
QY 292 ProLeuAspLeuIleValThrAlaLeuThrProH1sProProGlnIleGlnAlaThr 311
DB 834 CGTTATTAAGATTATCAAGACAGCGTTGACGCCCCCATCCAGCAAAACAGGCGAT 775
QY 312 GlyValThrLeuProThrSerValLeuPheIleAlaGlyHisAspThrAsnLeuAlaAsn 331
DB 774 GGTGTGACATTATCCACTTCAAGTGTCTGTTTATCCCGGACACGATACTATCTGGCAAT 715
QY 332 LeuGlyGlyAlaLeuGluLeuAsnTrpThrLeuProGlyGlnProAspAsnThrProPro 351
DB 714 CTGGCGCGGCGACCTGAGACTCAACTGACGCTTCCCGGTCAAGCCGATTAACGCGCGCA 655
QY 352 GlyGlyGluLeuValPheGluArgTrpArgArgLeuSerAspAsnSerGlnTrpIlegIn 371
DB 654 GGTGTGACATTATCCACTTCAAGTGTCTGTTTATCCCGGACACGATTAACGCGAGTTTCG 595
QY 372 ValSerLeuValPheGlnThrLeuGlnGlnMetArgAspIleThrProLeuSerLeuAsn 391
DB 594 GTTTCGCTGCTTCCACTTCAAGCAGATGCTGATTAACGCCGCTGCTATTAAT 535
QY 392 ThrProProGlyGlyGluValIleValLeuThrLeuAlaGlyCysGlnGluArgAsnAlaGlnGly 411
DB 534 ACGCGCGCGGAGAGGTGAACCTGACCTGCGCAGGATGTGAAGAGCGAAATGCCGAGGCG 475
QY 412 MetCysSerLeuAlaGlyPheThrGlnIleValAsnGluAlaArgIleProAlaCysSer 431
DB 474 ATGTTTGTGGCAGGTTTACCGCAATCTGTAATGAAGCAGCATACCGGCTGCAGT 415
QY 432 Leu 432
DB 414 TTG 412

RESULT 7
LOCUS CN754382/1 868 bp mRNA linear EST 19-MAY-2004
DEFINITION ID0AA12DE01RM1 ApMs Acyrthosiphon pisum cDNA clone ID0AA12DE01
5', mRNA sequence.
ACCESSION CN754382
VERSION CN754382.1 GI:47519379
KEYWORDS EST.
SOURCE Acyrthosiphon pisum (pea aphid)
ORGANISM Acyrthosiphon pisum
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.
1 (bases 1 to 868)
Hunter, W., Martinez-Torres, D., Rahbe, Y., Sabater-Munoz, B.,
Stern, D., Tagu, D. and Wincker, P.
An expressed sequence tags database for the pea aphid Acyrthosiphon
pisum
Unpublished (2004)
JOURNAL
COMMENT Contact: D. Tagu
INRA Rennes
UMR BIO3P BP 35327, F-35653 Le Rheu Cedex France
Tel: +33.2.23.48.51.65
Fax: +33.2.23.48.51.50
Risk of contamination by bacterial sequences from obligatory
(Buchnera) or facultative endosymbionts. These sequences were
obtained in the frame of the International Consortium of Aphid
Genomics in collaboration with Genoscope
PCR primers
FORWARD: CAGGAAACAGCTATGACC
Plate: 5 row: C column: 9.
Location/Qualifiers
1..895
/organism="Acyrthosiphon pisum"
/mol_type="mRNA"
/cultivar="developmentstage"
/db_xref="taxon:7029"
/clone="ID0AA5DC09"
/tissue_type="whole insect"
/dev_stage="nymphs and adults (parthenogenetic females)"
/lab_host="XLA Blue"
/clone_id="ApMs"
/note="Vector: pBS-SK minus, site 1: EcoRI, site 2: XhoI;
Sample name: ID0AAA ; Plant growth place: Department of

JOURNAL
COMMENT
pismum
Unpublished (2004)
Contact: D. Tagu
INRA Rennes
UMR BIO3P, BP 35327, F-35653 Le Rheu Cedex France
Tel: +33.2.23.48.51.65
Fax: +33.2.23.48.51.50
Risk of contamination by bacterial sequences from obligatory (Buchnera) or facultative endosymbionts. These sequences were obtained in the frame of the International Consortium of Aphid Genomics in collaboration with Genoscope

PCR Primers
FORWARD: CAGGAACAGCTATGACC
Plate: 12 row: E column: 1.
Location/Qualifiers
1. 868
/organism="Acyrthosiphon pisum"
/mol_type="mRNA"
/culturivar="developmentstage"
/db_xref="taxon:7029"
/clone="ID0AA1D801"
/tissue_type="whole insect"
/dev_stage="nymphs and adults (parthenogenetic females)"
/lab_host="X11-Blue"
/clone_lib="ApMS"
/note="Vector: pBS-SK minus; Site_1: EcoRI; Site_2: XhoI; Sample name: ID0AAA; Plant growth place: Department of Ecology & Evolutionary Biology, Princeton University; Soil conditions: Soil; Sowing date: 01/06/1999; Harvesting date: 01/06/1999; Stress date: no stress; Description: Aphids inoculated on one-week old Vicia faba under non-sterile conditions. All parthenogenetic stages and both winged and wingless adults were collected for library construction.; experimental condition: long photoperiod (16-hr light/8-hr dark at 18 c)"

ORIGIN

Alignment Scores:
Pred. No.: 5.37e-68 Length: 868
Score: 778.00 Matches: 151
Percent Similarity: 99.3% Conservative: 0
Best Local Similarity: 99.3% Mismatches: 1
Query Match: 34.5% Indels: 1
DB: 8 Gaps: 0

US-10-601-319-10 (1-432) x CN754382 (1-868)

Oy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleValThrAla 300
Db 866 CGCAGCGCAAGATT-GCCCGCAGCGCGCCGACCCCGTTATTTGATTCAGACGACG 808
Oy 301 LeuThrProHisProProGlnIleValArgValIleThrLeuProThrSerValIleu 320
Db 807 TTGACGCCCATCCACCCCAAAAACAGCGGTATGTGTGACATTCACCACTTCAGTGTG 748
Oy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeuGlnLeuAsnTrp 340
Db 747 TTTATCGCCGACAGATTAATCTGGCAATCTCGCGGCGACCTGAGCTCAACGCG 688
Oy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyLeuValPheGlnIleArgTrp 360
Db 687 ACGCTTCCCGGTACCGCGGATTAACGCGCGCGAGTGTGAATCTGTTTGAACGCTCG 628
Oy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 627 CGTGGCTTAAGCGATTAACGACGATTCAGGTTTCGTGTCTTCACAGCTTACAG 568
Oy 381 GlnMetArgAspIleValThrProLeuSerLeuAsnThrProProGlyGlyValIleValLeuThr 400
Db 567 CAGATGCGTGAATAAACCCGCTGTCATTAATACGCGCGGAGAGGTGAACCTGACG 508
Oy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420

Db 507 CTGGCAGATGTGTAAGACGAATGCGCAGGCGCATGTTGTTGTCAGGTTTACGCAA 448
Oy 421 IleValAsnGluValAlaArgIleProAlaCysSerLeu 432
Db 447 ATCGTGAATGAAGACGCGCATACCGGCTGTGCGATTG 412

RESULT 8
CN757565/c 864 bp mRNA linear EST 20-MAY-2004
LOCUS ID0AA1D099RM1 ApMS Acyrthosiphon pisum cDNA clone ID0AA1D09 5',
DEFINITION mRNA sequence.

ACCESSION CN757565
VERSION CN757565.1 GI:47531488
KEYWORDS EST.
SOURCE Acyrthosiphon pisum (pea aphid)

ORGANISM Acyrthosiphon pisum
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.

REFERENCE 1 (bases 1 to 864)
AUTHORS Hunter,W., Martinez-Torres,D., Rahbe,Y., Sabater-Munoz,B., Stern,D., Tagu,D. and Wincker,P.
TITLE An expressed sequence tags database for the pea aphid Acyrthosiphon pisum
JOURNAL Unpublished (2004)
COMMENT Contact: D. Tagu

INRA Rennes
UMR BIO3P, BP 35327, F-35653 Le Rheu Cedex France
Tel: +33.2.23.48.51.65
Fax: +33.2.23.48.51.50
Risk of contamination by bacterial sequences from obligatory (Buchnera) or facultative endosymbionts. These sequences were obtained in the frame of the International Consortium of Aphid Genomics in collaboration with Genoscope

PCR Primers
FORWARD: CAGGAACAGCTATGACC
Plate: 1 row: F column: 9.
Location/Qualifiers
1. 864
/organism="Acyrthosiphon pisum"
/mol_type="mRNA"
/culturivar="developmentstage"
/db_xref="taxon:7029"
/clone="ID0AA1D801"
/tissue_type="whole insect"
/dev_stage="nymphs and adults (parthenogenetic females)"
/lab_host="X11-Blue"
/clone_lib="ApMS"
/note="Vector: pBS-SK minus; Site_1: EcoRI; Site_2: XhoI; Sample name: ID0AAA; Plant growth place: Department of Ecology & Evolutionary Biology, Princeton University; Soil conditions: Soil; Sowing date: 01/06/1999; Harvesting date: 01/06/1999; Stress date: no stress; Description: Aphids inoculated on one-week old Vicia faba under non-sterile conditions. All parthenogenetic stages and both winged and wingless adults were collected for library construction.; experimental condition: long photoperiod (16-hr light/8-hr dark at 18 c)"

ORIGIN

Alignment Scores:
Pred. No.: 5.53e-67 Length: 864
Score: 768.00 Matches: 150
Percent Similarity: 99.3% Conservative: 0
Best Local Similarity: 99.3% Mismatches: 1
Query Match: 34.0% Indels: 1
DB: 8 Gaps: 0

US-10-601-319-10 (1-432) x CN757565 (1-864)

Oy 282 ThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleValThrAlaLeu 301
Db 863 ACGCAGAGGTTGCCCGCAGCGCGCCGCGCTTATTATGATTC-ATCAAGACAGCGTTG 805

QY 302 ThrProHisProProGlnIysGlnAlaIyGlyValThrLeuProThrSerValLeuPhe 321
 Db 804 AGCCCCCATTCACCGGAAAAACAGCGCTATGCTGTAACATTACCTTCAAGCTGCTTT 745
 QY 322 ILeaAGLYHIAAPrThrAsnLeuAlaAsnLeuGlyGlyValAlaGlnIleuAsnTPThr 341
 Db 744 ATCGCGGACACGATCTATCTGCGAAATCTCGCGGCGGACCTGAGCTCAACTGAGG 685
 QY 342 LeuProGlyGlnProAspAsnThrProProGlyGlyGlyIleuValPheGlnArgTPArg 361
 Db 684 CTTCCTCGGTACGCGGATTAACACGCGCGAGGTGTGAACTGGTGTGAAACGCTGGCGT 625
 QY 362 ArgLeuSerAspAsnSerGlnTPrlIleGlnValSerLeuValPheGlnThrLeuGlnGln 381
 Db 624 CGGCTACGCAATACAGCCAGTGAGATTACAGTTTCCTGCTCTTCAAGCTTACAGCAG 565
 QY 382 MetArgAspIyThrProLeuSerLeuAsnThrProProGlyGlyValIySerLeuThrLeu 401
 Db 564 ATGCGGTATTAACCGCGCTGTCTATTAAATACGCCGCCCGGAGAGGTGAACTGACCTG 505
 QY 402 AlaGlyCyseGlnIyArgAsnAlaGlnIyMetCyseSerLeuAlaGlyPheThrGlnIle 421
 Db 504 GCAGGATGTGAAGAGCGAAATGCGCAGGCGCATGTGTTGCTGGCAGGTTTACGCAATC 445
 QY 422 ValAsnGluAlaArgIleProAlaCyseSerLeu 432
 Db 444 GTGAATGAAGACGCACTACCGCGGTGCAAGTTTG 412

RESULT 9 383 bp mRNA linear EST 18-MAY-2001
 LOCUS AM036132
 DEFINITION EST1774508 tomato seed, TAMU Lycopersicon esculentum cDNA clone
 CLEB1E23 similar to periplasmic phosphoanhydride phosphohydrolase
 precursor, putative, mRNA sequence.

AM036132
 AM036132.1 GI:5894811
 EST.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Lycopersicon esculentum (Solanum lycopersicum)

Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 383)

REFERENCE

AUTHORS

Alcala,J., Vrebalov,J., White,R., Matera,A.L., Vision,T.,
 Holt,I.E., Liang,F., Upton,J., Romling,C.M., Craven,M.B.,
 Fujii,C.Y., Bowman,C.L., Nierman,W., Frazer,C.M., Venter,J.C.,
 Martin,G.B., Tanksley,S.D. and Giovannoni,J.
 Generation of ESTs from tomato seed tissue
 Unpublished (1999)

TITLE

JOURNAL

COMMENT
 Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>

3 prime sequence.
 Location/Qualifiers

FEATURES

source

1..383
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="CLEB1E23"
 /tissue_type="seeds"
 /dev_stage="quiescent seed"
 /lab_host="XLI-Blue MRF"
 /clone_lib="tomato seed, TAMU"
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
 XhoI; cLBR - Tomato seed EST library. Directionally cloned
 cDNAs inserted into pBluescript SK(-) at 5' end with
 EcoRI and 3' end with XhoI site."

ORIGIN

Alignment Scores:

Pred. No.: 8,57e-53 Length: 383
 Score: 623.00 Matches: 124
 Percent Similarity: 97.6% Conservative: 0
 Best Local Similarity: 97.6% Mismatches: 3
 Query Match: 27.6% Indels: 0
 DB: 7 Gaps: 0

US-10-601-319-10 (1-432) x AM036132 (1-383)

QY 117 LysThrGlyGluAlaPheAlaAlaGlyLeuAlaProAspCyseAlaIleThrValHisThr 136
 Db 2 AAAACAGCGGAGGACCTTCCGCCCGCGCTGACCTGATGTGCAATACCGTACATACC 61
 QY 137 GlnAlaAspThrSerSerProAspProLeuPheAsnProLeuIyThrGlyValCyseGln 156
 Db 62 CAGCAGATACCTCCAGTCCCATCCGATCTGATTTAATCCCTAAAACTGGCTTGGCAA 121
 QY 157 LeuAspAsnAlaAsnValThrAspAlaIleLeuGluArgAlaGlyIySerIleAlaAsp 176
 Db 122 CTGATTAACGCGAAGCTGACTGACCGATCTCAGCAGGCGAGAGGTCAATTGCTGAC 181
 QY 177 PheThrGlyHisIyTrGlnThrAlaPheArgGlyLeuGluArgValIleuAsnPheProGln 196
 Db 182 TTTACCGGCAATCGGCAACGCGCTTCCGCAACTGGAACGGGTCTTAATTTCCGCCAA 241
 QY 197 SerAsnLeuCyseLeuIyArgGlyIyIySerIyAspGlyIySerCyseSerLeuThrGlnAlaLeu 216
 Db 242 TCAACTGTGTGCTTAAAGCTGAGAAACAGACGAAAGCTTCAATTACGACGAGCATTA 301
 QY 217 ProSerGlyLeuIyValSerIleAspCyseValSerLeuThrGlyAlaValSerLeuAla 236
 Db 302 CCATCGAATCACTAAGGTAGCGCGCAATGCTCATTAACGGTGGTAAAGCTTCGCA 361
 QY 237 SerMetLeuThrGluIlePhe 243
 Db 362 TCAATGCTGACGAGATATTT 382

RESULT 10 383 bp mRNA linear EST 18-MAY-2001
 LOCUS AM036134/C
 DEFINITION EST1774510 tomato seed, TAMU Lycopersicon esculentum cDNA clone
 CLEB1E23 similar to periplasmic phosphoanhydride phosphohydrolase
 precursor, putative, mRNA sequence.

AM036134
 AM036134.1 GI:5894813
 EST.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Lycopersicon esculentum (Solanum lycopersicum)

Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 383)

REFERENCE

AUTHORS

Alcala,J., Vrebalov,J., White,R., Matera,A.L., Vision,T.,
 Holt,I.E., Liang,F., Upton,J., Romling,C.M., Craven,M.B.,
 Fujii,C.Y., Bowman,C.L., Nierman,W., Frazer,C.M., Venter,J.C.,
 Martin,G.B., Tanksley,S.D. and Giovannoni,J.
 Generation of ESTs from tomato seed tissue
 Unpublished (1999)

TITLE

JOURNAL

COMMENT
 Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.
 Location/Qualifiers

FEATURES

source

1..383
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="CLEB1E23"
 /tissue_type="seeds"

/dev stage="quiescent seed"
/lab_host="XII-Blue MRP."
/clone_lib="tomato seed, TAMU"
/note="Vector: pbluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; cDNA - Tomato Seed EST library. Directionally cloned
cDNAs inserted into pbluescript SK(-) at 5' end with
EcoRI and 3' end with XhoI site."

ORIGIN

Alignment Scores:

Pred. No.:	8,57e-53	Length:	383
Score:	623.00	Matches:	124
Percent Similarity:	97.6%	Conservative:	0
Best Local Similarity:	97.6%	Mismatches:	3
Query Match:	27.6%	Indels:	0
DB:	7	Gaps:	0

US-10-601-319-10 (1-432) x AM036134 (1-383)

Qy 117 LysThrglygluaIaphaAlaIaGlyLeuAlaProaPpCyAlaIleThrValHisThr 136
Db 382 AAAACAGCGCAAGCTTCCGCGCGCTGGCACTGACTGCAATACCTACATACC 323
Qy 137 GlnAlaapThrSerSerProaPpProleuPheanProleuYthrGlyValCysGln 156
Db 322 CAGCAGATACGTCAGTCCGATCCGATCCGTTATTTATCTTAAATAAAGTGGCTTGC 263
Qy 157 LeuapPanaAlaenValThraPAlaIleLeuGluAlaGlyGlySerIleAlaAsp 176
Db 262 CTGATTAACGGGAACGTACTGACGCGATCCTCGACAGGAGGATCAATGGCTGAC 203
Qy 177 PheThrglyHisYrGlnThraIaphaArgIuleuGluArgValIleuAsnPhaProGln 196
Db 202 TTTACCGGCAATCGGCAACCGCTTTCGCAACTGGAACCGGTCTTAATTTCCGCA 143
Qy 197 SeranLeuCyLeuYsArgIuLysGlnaPpGlnuSerYsSerIleuThraIaLeu 216
Db 142 TCAAACTTGCTCTTAAAGTGAAGAACAGACGAAAGCTGTTCATTACGACGATTA 83
Qy 217 ProSerGluLeuYsValSerAlaAspCysValSerIleuThrglyAlaValSerIleuAla 236
Db 82 CCATCGGAACCTCAAGGTAGCGCCGACATGTCTCATTTAACCGGTGCGTAAAGCTCGCA 23
Qy 237 SerMetLeuThrglyIlePhe 243
Db 22 TCATGCTGACGAGATTTT 2

RESULT 11 BE520240 393 bp mRNA linear EST 19-MAR-2001
LOCUS M1B12STM Arabidopsis developing seed Arabidopsis thaliana cDNA
DEFINITION clone M1B12 5', mRNA sequence.
ACCESSION BE520240
VERSION BE520240.1 GI:9778242
KEYWORDS EST.
SOURCE Arabidopsis thaliana (chale crese)
ORGANISM Arabidopsis thaliana

REFERENCE 1 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 393)
AUTHORS White, J.A., Todd, J., Newman, T., Focks, N., Girke, T., Martinez de
Larduyva, O., Javorecki, J.G., Ohlrogge, J., and Benning, C.
TITLE A new set of Arabidopsis expressed sequence tags from developing
seeds. The metabolic pathway from carbohydrates to seed oil
JOURNAL Plant Physiol. 124 (4), 1582-1594 (2000)
PUBMED 1115876

COMMENT Contact: Benning, C
Dept. of Biochemistry & Molecular Biology
Michigan State University
224 Biochemistry, Michigan State University, East Lansing, MI
48824, USA
Tel: 517 355 1609

Fax: 517 353 9334
Email: benning@msu.edu
Michigan State University DNA Sequencing Facility Arabidopsis
Biological Resource Center, The Ohio State University, 309 Bocrany &
Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210 USA, FAX:
6142920603 TEL: 6142929371.
location/Qualifiers
1..393
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="M1B12"
/rname type="seed"
/rstage="5-13 days after flowering"
/dev stage="E.coli"
/clone_lib="Arabidopsis developing seed"
/note="Organ: Developing seed; Vector: pbluescript SK-;
Site_1: EcoRI; Site_2: XhoI"

FEATURES

source

ORIGIN

Alignment Scores:

Pred. No.:	1.42e-52	Length:	393
Score:	621.00	Matches:	124
Percent Similarity:	95.4%	Conservative:	1
Best Local Similarity:	94.7%	Mismatches:	6
Query Match:	27.5%	Indels:	0
DB:	7	Gaps:	0

US-10-601-319-10 (1-432) x BE520240 (1-393)

Qy 107 IleIleAlaaspValaIaPpGluArgThraYrGlnThrglyGlnAlaPheAlaIle 126
Db 393 ATTTTCTGATGTGACGAGCGTACCCGCTAANAACGAGCTTCCGCCGGGNTG 334
Qy 127 AlaProaPpCyAlaIleThraValHisThrGlnAlaAspThrSerSerProaPpProleu 146
Db 333 GCACCTGACGTGCAATTAACGATACATACCAAGCAGATACGTCAGTCCGTTA 274
Qy 147 PheanProleuYthrGlyValCysGlnLeuAspPanaIleValThraPAlaIle 166
Db 273 TTTAATCTCTTAAAAACCTGCGTTTGCACTGATAACCGCAAGTGAAGCGATC 214
Qy 167 LeuGluArgAlaGlySerIleAlaAspPheThrglyHisYrGlnThraIaphaArg 186
Db 213 CTCAGCAGGCGAGAGGCTCAATTGCTGACTTTACCGGCAATCGGCAAGCGCTTCCG 154
Qy 187 GluLeuGluArgValLeuAsnPhaProGlnSeranLeuCyLeuYsArgIuLysGln 206
Db 153 GAACGTGGAACGGGTGCTTAATTTCCGCAATCAAACTTGCTTAAACGTGAAGAACAG 94
Qy 207 AspGlnuSerYsSerIleuThrglyAlaLeuProSerGluLeuYsValSerAlaAspCys 226
Db 93 GACGAAAGCTGTTCATTAAACGAGGCAATTCATCGGAACCTCAAGGTGAGCGCGACAAAT 34
Qy 227 ValSerLeuThrglyAlaValSerIleAlaSer 237
Db 33 GTCTCATTAAGCGCTGCGGTAAAGCTTGCATCA 1

RESULT 12 BU713769 531 bp mRNA linear EST 23-OCT-2003
LOCUS SJAABH01 Adult S1C 7/94 Schistosoma japonicum cDNA similar to
DEFINITION bp|P0102|PRA_BO01 PERIPLASMIC APPA PROTEIN PRECURSOR (LINLUNES:
PHOSPHOANHYDRIDE PHOSPHOHYDROLASE (PH 2.5 ACID PHOSPHATASE) (AP);
6-PHATASE 1, mRNA sequence.
ACCESSION BU713769
VERSION BU713769.1 GI:28321125
KEYWORDS EST.
SOURCE Schistosoma japonicum
ORGANISM Schistosoma japonicum
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
Strigeida; Schistosomatidae; Schistosomatidae; Schistosoma.

REFERENCE
1 (bases 1 to 531)
AUTHORS
Hu, W., Yan, O., Shen, D. K., Liu, F., Zhu, Z. D., Song, H. D., Xu, X. R., Wang, Z. J., Rong, Y. P., Zeng, L. C., Wu, J., Zhang, X., Wang, J. J., Xu, X. N., Wang, S. Y., Fu, G., Zhang, X. L., Wang, Z. Q., Brindley, P. J., McManus, D. P., Xue, C. L., Peng, Z., Chen, Z., and Han, Z. G.
Evolutionary and biomedical implications of a Schistosoma japonicum complementary DNA resource
Nat. Genet. 35 (2), 139-147 (2003)
JOURNAL
12973349
COMMENT
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn.
Location/Qualifiers
FEATURES
source
1..531
/organism="Schistosoma japonicum"
/mol_type="mRNA"
/strain="Chinese (Anhui) strain"
/db_xref="taxon:6182"
/sex="Male and female"
/tissue_type="Whole body"
/dev_stage="Adult worms"
/lab_host="Mouse and rabbit"
/clone_1lb="Adult 8yc 7/94"
/note="Vector: Lambda ZAP-11 XR.; Site 1: EcoR I; Site 2: XhoI I; Several hundred adult Schistosoma japonicum (Anhui, P. R. China, strain), of mixed sex, were perfused from the mesenteries of experimentally infected mice and rabbits at the Queensland Institute of Medical Research, Brisbane, Australia (QIMR), and stored for several months in liquid nitrogen. Subsequently, mRNA was isolated at the QIMR from lysates of these worms by oligo dt chromatography, using a kit from Pharmacia. The mRNA was then shipped to Clontech, Palo Alto, CA, USA, who constructed a cDNA library. First strand synthesis was primed with an oligo-dt-XhoI-primer and synthesized using M-MuLV reverse transcriptase. Second strand synthesis was accomplished with RNase H and T4 DNA polymerase. The double stranded cDNA was ligated to EcoRI linkers, digested with EcoRI and XhoI, and ligated into the phagemid vector lambda ZAP II XR. After construction of this directional library by Clontech, it was returned to the QIMR. During analysis of the library at the QIMR, we have found that a small percentage, 2% to 3%, of the clones contain inserts that appear to be highly homologous to sequences from salmonid fishes, as determined by homology comparisons using BLAST and by Southern hybridization analysis to genomic DNA from salmon (Sigma Chemical Co., St. Louis, MO) under stringent washing conditions. The remainder of the clones appear to contain S. japonicum sequences."

ORIGIN
Alignment Scores:
Fred. No.: 1.46e-48 Length: 531
Score: 583.50 Matches: 133
Percent Similarity: 79.5% Conservative: 7
Best Local Similarity: 75.6% Mismatches: 30
Query Match: 25.8% Indels: 7
DB: 3 Gaps: 1

US-10-601-319-10 (1-432) x BU713769 (1-531)

QY 224 A1AaPpCySvAlSerIeuThrG1yAlaValSerLeuAlaSerMerIeuThrG1uIlePhe 243
DB 1 GCCGCAATGTCTTAAACCGGTGGGTAACCTCGCATCAATGACGAGATGGGG 60
QY 244 LeuIeuGInGInAlGInG1yMeCProG1uProG1yTTPG1yArg1IeThzApSerHis 263
DB 61 CTCCTGCAACAAGCAGGGAATGCCGAGCGGGGTGGGGAAGATACCGATTCAAC 120

QY 264 G1TTPAeThrThrLeuIeuSerIeuHiSaAna1aGInPheAspLeuIeuGInAArgThrPro 283
DB 121 CAGTGGAAACCTTGATTCGATTTGATGATTAACGGGCAATTTATTTGGCTACAAAGCGCCA 180
QY 284 GIUValAlAaRySerAryglAthrProIeuLeuAapLeu1IleTyThra1A1euThrPro 303
DB 181 GAGGTTGCCGCAAGCGCGCCACCCCGATTAAGATTGATACAGACAGCGTTGACGCC 240
QY 304 H1sProProGInLyGInAlaTyG1yValThrIeuProThrSexValIeu-Phe1IleA1 323
DB 241 CGTTATATCATATTATGATTTGCAAGTGGAGAGACAATTCCTCTTGAGTACGAAAAAAG 300
QY 323 aG1yHiSaPThrAaenLeuAlaAsnLeuG1yG1yAlaIeuG1uIeuAenThrThrIeuDr 343
DB 301 CGGCAAGCGAGCTGATGATTAACCTCGCGCCGCACTGAGCTCAACTGACCGCTGTC 360
QY 343 oG1yGInProAaPaenThrProProG1yG1yG1uIeu-ValPheG1uArgTrp-ArgArg 362
DB 361 CAGGCAAGCCGATTAACAGCCCGCAGAGCGGAAGCTGAGTGTTAACGCTGGCGCTCG 420
QY 363 LeuSerAaPaenSerGInThrP1IleGInValSerLeuValPheGInThrIeuGInIleMet 382
DB 421 CTAACGATTAACAGCAAGTGAATTCAGGTTTGGCTGCTTCCAGACTTA-CAGCAGATG 479
QY 383 ArgAaP1yThrProIeuSerIeuAaThrProProG1yG1u 396
DB 480 CCGCGGGGGCTTCCTTGGGG-----CCTTATGGGGAT 512

RESULT 13
LOCUS CN759004/c
DEFINITION ID0AA24BC04RM1 ApMS Acyrthosiphon pisum cDNA clone ID0AA24BC04
5', mRNA sequence.
ACCESSION CN759004
VERSION CN759004.1 GI:47532927
KEYWORDS EST
SOURCE Acyrthosiphon pisum (pea aphid)
ORGANISM Acyrthosiphon pisum
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes; Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.
REFERENCE 1 (bases 1 to 706)
Huner, W., Martinez-Torres, D., Rahbe, Y., Sabater-Munoz, B., Stern, D., Tagu, D., and Winkler, P.
An expressed sequence tags database for the pea aphid Acyrthosiphon pisum
Unpublished (2004)
JOURNAL
COMMENT
Contact: D. Tagu
INRA Rennes
UMR BIO3P, BP 35327, F-35653 Le Rheu Cedex France
Tel: +33-2-23-48-51-65
Fax: +33-2-23-48-51-50
Risk of contamination by bacterial sequences from obligatory (Buchnera) or facultative endosymbionts. These sequences were obtained in the frame of the International Consortium of Aphid Genomics in collaboration with Genoscope
PCR Primers
FORWARD: CAGGAACAGCTATGACC
Plate: 24 row: C column: 4.
FEATURES
source
1..706
/organism="Acyrthosiphon pisum"
/mol_type="mRNA"
/cultiVar="developmentstage"
/db_xref="taxon:7029"
/clone="ID0AA24BC04"
/tissue_type="whole insect"
/dev_stage="nymphs and adults (parthenogenetic females)"
/lab_host="X11-Blue"
/clone_1lb="APMS"
/note="Vector: pBS-SK minus; Site 1: EcoRI; Site 2: XhoI;
Sample name: ID0AAA ; Plant growth place: Department of

Ecology & Evolutionary Biology, Princeton University ;
Soil conditions: Soil ; Sowing date: 01/06/1999 ;
Harvesting date: 01/06/1999 ; Stress date: no stress ;
Description: Aphids inoculated on one-week old *Vicia faba*
under non-sterile conditions. All parthenogenetic stages
and both winged and wingless adults were collected for
library construction. ; experimental condition: long
photoperiod (16-hr light/8-hr dark at 18 c°)

ORIGIN

Alignment Scores:

Pred. No.:	4.02e-42	Length:	706
Score:	522.00	Matches:	98
Percent Similarity:	100.0%	Conservative:	0
Best local Similarity:	100.0%	Mismatches:	0
Query Match:	23.1%	Indels:	0
DB:	8	Gaps:	0

US-10-601-319-10 (1-432) x CN759004 (1-706)

Qy 335 AAlaenGluLeuAaNTPrThleuProGlyGlnProAspAaNTPrProGlyGlyGlu 354
Db 705 GCACTGAGCTCAACTGACCGCTCCCGTACGCCGATTAACGCCCGCAGTGTAA 646
Qy 355 LeuValPheGluArgTrpArgArgLeuSerAspAsnSerGlnTrpLleGlnValSerLeu 374
Db 645 CTGGGTGTTGAACGCTGGCGCTGGCTAAGCGATTAACACCGATTCAGGTTTCGCTG 586
Qy 375 ValPheGlnTrhLeuGlnGlnMetArgAspLysThrProLeuSerLeuAaNTPrProPro 394
Db 585 GTCTTCCAGACTTTACACAGACATGCGTATAAACGCCGCTGTCAATTAATACGCCGCC 526
Qy 395 GlyGluValLysLeuThrLeuAlaGlyCysGlnGluGlyGlnAlaGlnGlyMetCysSer 414
Db 525 GGAAGGGGAACGACCTCGACGATGTGAAGGCAAAATGGCAGGCGCATGTGTG 466
Qy 415 LeuAlaGlyPheThrGlnLleValAsnGluAlaArgLleProAlaCysSerLeu 432
Db 465 TTGGCAGGTTTACGCAATCGTGAATGAACGACCAATACCGCGCTGCAGTTTG 412

RESULT 14
AA545747/c 354 bp mRNA linear EST 12-MAY-1999
LOCUS HBMSF1B4-REV Human Bone Marrow Stromal Fibroblast Homo sapiens cDNA
DEFINITION clone HBMSF1B4 5', mRNA sequence.
ACCESSION AA545747
VERSION AA545747.1 GI:2307026
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 354)

REFERENCE
AUTHORS Jia,L., Robey,P., Young,M., Schuler,G., Powell,J., Yang,L.,
Lemon,G., Hillier,L., Allen,M., Bowles,L., Geisel,S., Kucaba,T.,
Marrin,M., Martin,J., Stepien,M., Tan,F., Theising,B., Bowers,Y.,
Wylie,T., Waterston,R., Wilson,R. and Francomano,C.
TITLE WashU-MGB/MGRI EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Libin Jia
National Human Genome Research Institute
10/10c101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
Tel: 301-402-4877
Fax: 301-496-7157

FEATURES
source
1..354
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

/clone="HBMSF1B4"
/sex="Male and Female"
/tissue_type="bone marrow"
/cell_type="stromal fibroblast"
/dev_stage="mixed"
/lab_host="XLI-Blue"
/clone_lib="Human Bone Marrow Stromal Fibroblast"
/notes="Vector: pBluescript; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Alignment Scores:

Pred. No.:	1.28e-38	Length:	354
Score:	483.00	Matches:	108
Percent Similarity:	91.6%	Conservative:	1
Best local Similarity:	90.8%	Mismatches:	9
Query Match:	21.4%	Indels:	4
DB:	1	Gaps:	0

US-10-601-319-10 (1-432) x AA545747 (1-354)

Qy 166 ILleuGluValArgAlaGlySerLleAlaAspPheThrGlyHisTyrglnThr-Alaph 185
Db 352 GTCTTACGACGAGGACGAGGTCA-ATTGCTGA-TTTACGGGCAATCGGCAACGGGCTT 295
Qy 185 eArgGluLeuGluValArgValLeuAsnPheProGlnSerAsnLeuCysLeuysArgGlu 205
Db 294 TCGGAACTGGAACGGGTGTT-AATTTCCGCAATCAAACTTGTCCTTAACGTGAA 236
Qy 205 eGlnAspGlySerCysSerLeuThrGlnAlaLeuProSerGluLeuLysValSerAlaAs 225
Db 235 ACAGAGGAAGAGCTGTTCAATTAACGCGCATTAACCATCGAACTCAAGGTGAGCGCGA 176
Qy 225 pCysValSerLeuThrGlyValAlaValSerLeuAlaSerMetLeuThrGlnLlePheLeu 245
Db 175 CAATGTTCATTAACCGGTGGGTAGCCCTGCATCATGCTAGCGAGATATTTCCT 116
Qy 245 uGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgLleThrAspSerHisGlnTr 265
Db 115 GCAACAGACACAGGAGATGGCGAGCGGGGTGGGAGAGTACCGATTCACACCGAGTG 56
Qy 265 pAenThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGlnArgThrPro 283
Db 55 GAACACCTTGCTAATGTTGCATTAACGCCCAATTTATTGCTACACGACGCCA 1

RESULT 15
B619443 707 bp mRNA linear EST 01-OCT-2003
LOCUS B619443 NIBB Mochii normalized Xenopus early gastrula library
DEFINITION Xenopus laevis cDNA clone XLI89424 5', mRNA sequence.
ACCESSION B619443
VERSION B619443.1 GI:37258203
KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodidae; Xenopus; Xenopus.
1 (bases 1 to 707)

REFERENCE
AUTHORS Kiteyama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and
Kohara,Y.
TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasi Shin-i
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855

FEATURES
source
1..707
Location/Qualifiers
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:9606"

The information of this clone is available through the following
URL:
http://xenopus.nibb.ac.jp.
Location/Qualifiers

source

```
1..707
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="XL189d24"
/tissue_type="whole embryo"
/dev_stage="stage 10.5"
/clone_lib="NIBB Mochiz normalized Xenopus early gastrula
library"
```

ORIGIN

Alignment Scores:

Pred. No.:	4,29e-34	Length:	707
Score:	443.00	Matches:	84
Percent Similarity:	98.8%	Conservative:	0
Best Local Similarity:	98.8%	Mismatches:	1
Query Match:	19.6%	Indels:	0
DB:	2	Gaps:	0

US-10-601-319-10 (1-432) x BU619443 (1-707)

```
QY      251 MetProGluProGlyTTrpGlyArgIleThrAspSerHisGlnTrpAsnThrLeuSer 270
          |||||
Db       2  ATGCCGAGACCGGGGTGGGAGAGATCACCATTCACACAGTGGAAACCTTGCTAACT 61
          |||||
QY      271 LeuHisAsnAlaGlnPheAspLeuGlnArgThrProGluValAlaArgSerArgAla 290
          |||||
Db       62 TTGCATTAACGGCCCAATTATTATTGCTACAAACGACGCCAGAGTTGCCCGACGCCGCC 121
          |||||
QY      291 ThrProLeuLeuAspLeuIleLysThrAlaLeuThrProHisProProGlnLysGlnAla 310
          |||||
Db      122 ACCCGATTATTCATTTGATTCAGACAGCGTTGACGCCCCCATCCACCGCAAAACAGGCG 181
          |||||
QY      311 TyrGlyValThrLeuProThrSerValLeuPheIleAlaGlyHisAspThrAsnLeuAla 330
          |||||
Db      182 TATGGTGTGACATTAACCACTTCAGTGTGTTTATCGCCGACACGATACTAATCTGGCA 241
          |||||
QY      331 AsnLeuGlyGlyAla 335
          |||||
Db      242 AATCTCGGGGGCGCA 256
          |||||
```

Search completed: June 14, 2006, 15:26:34
Job time : 4899.67 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2006, 10:17:34 ; Search time 189.248 Seconds

(without alignments)
211.554 Million cell updates/sec

Title: US-10-601-319-10

Perfect score: 2258
Sequence: 1 MKALIFPLSLIPITPSA.....CSLAGFTQIVNEARIPACSL 432

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: uniprot_sprot:*
2: uniprot_crembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query length	ID	Description
1	2182	96.6	432	1 PPA_ECOLI	P07102 escherichia
2	2182	96.6	432	2 Q23E1_SHISS	Q23E1 shigella so
3	2178	96.5	432	2 Q6GN88_ECOLI	Q6GN88 escherichia
4	2176	96.4	432	2 Q6RND6_ECOLI	Q6RND6 escherichia
5	2176	96.4	432	2 Q6RND7_ECOLI	Q6RND7 escherichia
6	2176	96.4	432	2 Q6RND8_ECOLI	Q6RND8 escherichia
7	2176	96.4	432	2 Q6RND9_ECOLI	Q6RND9 escherichia
8	2175	96.3	432	2 Q6RND9_ECOLI	Q6RND9 escherichia
9	2172	96.2	432	2 Q6RND9_ECOLI	Q6RND9 escherichia
10	2154	95.4	432	2 Q31YP3_SHIBS	Q31YP3 shigella bo
11	2149	95.2	432	2 Q6RND9_ECOLI	Q6RND9 escherichia
12	2147	95.1	432	2 Q6RND9_ECOLI	Q6RND9 escherichia
13	2145	95.0	432	2 Q6RND9_ECOLI	Q6RND9 escherichia
14	2144	95.0	432	2 Q6RND9_ECOLI	Q6RND9 escherichia
15	2144	95.0	432	2 Q6RND9_ECOLI	Q6RND9 escherichia
16	2143	94.9	432	2 Q6RND9_ECOLI	Q6RND9 escherichia
17	2130	94.3	432	2 Q6RND9_ECOLI	Q6RND9 escherichia
18	1370.5	60.7	432	2 Q6RND9_ECOLI	Q6RND9 escherichia
19	1357.5	60.1	432	2 Q6RND9_ECOLI	Q6RND9 escherichia
20	1087	48.1	444	2 Q6U677_9ENTR	Q6U677 obesumbacte
21	1063.5	47.1	444	2 Q6U677_9ENTR	Q6U677 obesumbacte
22	959.5	42.5	441	2 Q669R3_YERPS	Q669R3 yersinia ps
23	959.5	42.5	441	2 Q669R3_YERPS	Q669R3 yersinia ps
24	560	24.8	413	2 Q6RND9_ECOLI	Q6RND9 escherichia
25	556	24.6	413	2 Q6RND9_ECOLI	Q6RND9 escherichia
26	554	24.5	413	2 Q6RND9_ECOLI	Q6RND9 escherichia
27	553	24.5	413	2 Q6RND9_ECOLI	Q6RND9 escherichia
28	553	24.5	413	2 Q6RND9_ECOLI	Q6RND9 escherichia
29	552	24.4	413	2 Q6RND9_ECOLI	Q6RND9 escherichia
30	552	24.4	413	2 Q6RND9_ECOLI	Q6RND9 escherichia
31	547	24.2	413	2 Q6RND9_ECOLI	Q6RND9 escherichia

32	547	24.2	413	2	Q6RND9_ECOLI	Q6RND9 escherichia
33	536	23.8	417	1	Q6RND9_ECOLI	Q6RND9 escherichia
34	536	23.7	413	2	Q6RND9_ECOLI	Q6RND9 escherichia
35	536	23.7	414	2	Q6RND9_ECOLI	Q6RND9 escherichia
36	515	22.8	392	2	Q6RND9_ECOLI	Q6RND9 escherichia
37	512	22.7	443	2	Q6RND9_ECOLI	Q6RND9 escherichia
38	512	22.7	443	2	Q6RND9_ECOLI	Q6RND9 escherichia
39	491	21.7	515	2	Q6RND9_ECOLI	Q6RND9 escherichia
40	485	21.5	532	2	Q6RND9_ECOLI	Q6RND9 escherichia
41	477.5	21.1	435	2	Q6RND9_ECOLI	Q6RND9 escherichia
42	475.5	21.1	435	2	Q6RND9_ECOLI	Q6RND9 escherichia
43	471.5	20.9	433	2	Q6RND9_ECOLI	Q6RND9 escherichia
44	446	19.8	428	2	Q6RND9_ECOLI	Q6RND9 escherichia
45	436	19.3	368	2	Q6RND9_ECOLI	Q6RND9 escherichia

ALIGNMENTS

RESULT 1
PPA_ECOLI
ID PPA_ECOLI STANDARD; PRT; 432 AA.
AC P07102;
DT 01-APR-1988, integrated into UniProtKB/Swiss-Prot.
DT 01-AUG-1991, sequence version 2.
DT 07-MAR-2006, entry version 63.
DE Periplasmic appa protein precursor [includes: Phosphoanhydride
phosphohydrolyase (EC 3.1.3.2) (pH 2.5 acid phosphatase) (AP), 4-
phylase (EC 3.1.3.26)].
GN Name: appa; Ordered locus names: b0980;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND PROTEIN SEQUENCE OF 23-29.
RC STRAIN=K12;
RX MEDLINE=90368616; PubMed=2168385;
RA Dasea J., March C., Boquet P.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appa
reveals significant homology between pH 2.5 acid phosphatase and
glucose-1-phosphatase."
RT J. Bacteriol. 172:5497-5500(1990).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RT Science 277:1453-1474(1997).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=K12 / W3110;
RX MEDLINE=97061202; PubMed=8905232; DOI=10.1093/dnares/3.3.137;
RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
Kishida S., Kitagawa M., Makino K., Masuda S., Miki T., Mizouchi K.,
Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
Yano M., Horikuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 12.7-28.0 min region on the linkage map."
RT DNA Res. 3:117-155(1996).
RN [4]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 1-112.
RX MEDLINE=87271766; PubMed=3038201; DOI=10.1016/0304-8771(90)045-9;
RA Tonati E., Danchin A.;
RT "The structure of the promoter and amino terminal region of the pH 2.5
acid phosphatase structural gene (appa) of E. coli: a negative control
of transcription mediated by cyclic AMP.";

RL Blochimie 69:215-221(1987).
 RN [5]
 RP NUCLEOTIDE SEQUENCE [GENOMIC DNA] OF 1-17.
 RC STRAIN=K12;
 RX MEDLINE=92049231; PubMed=1658595; DOI=10.1007/BF00267454;
 RA Dassa J., Feisli H., Marck C., Dion M., Kieffer-Bontemps M.,
 RA Boquet P.L.;
 RT "A new oxygen-regulated operon in *Escherichia coli* comprises the genes
 RT for a putative third cytochrome oxidase and for pH 2.5 acid
 RT phosphatase (appa).";
 RL Mol. Genet. 229:341-352(1991).
 RN [6]
 RP CHARACTERIZATION, AND PROTEIN SEQUENCE OF 23-34.
 RA Greiner R., Jany K.-D.;
 RT "Characterization of a phytase from *Escherichia coli*.";
 RL Biol. Chem. Hoppe-Seyler 372:664-665(1991).
 RN [7]
 RP CHARACTERIZATION, AND PROTEIN SEQUENCE OF 23-35.
 RX MEDLINE=93256556; PubMed=8387749; DOI=10.1006/abbi.1993.1261;
 RA Greiner R., Komletzky U., Jany K.-D.;
 RT "Purification and characterization of two phytases from *Escherichia coli*.";
 RL Arch. Biochem. Biophys. 303:107-113(1993).
 RN [8]
 RP CHARACTERIZATION.
 RX MEDLINE=20161462; PubMed=10696472; DOI=10.1139/cjm-46-1-59;
 RA Golovan S., Wang G., Zhang J., Foreberg C.W.;
 RT "Characterization and overproduction of the *Escherichia coli* appa
 RT encoded bifunctional enzyme that exhibits both phytase and acid
 RT phosphatase activities.";
 RL Can. J. Microbiol. 46:59-71(2000).
 RN [9]
 RP MUTAGENESIS.
 RX MEDLINE=93054596; PubMed=1429631;
 RA Ostalan K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.-M.,
 RA van Eeten R.L.;
 RT "Overexpression, site-directed mutagenesis, and mechanism of
 RT *Escherichia coli* acid phosphatase.";
 RL J. Biol. Chem. 267:22830-22836(1992).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (2.05 ANGSTROMS).
 RX MEDLINE=20122624; PubMed=10655611; DOI=10.1038/72371;
 RA Lim D., Golovan S., Foreberg C.W., Jia Z.;
 RT "Crystal structures of *Escherichia coli* phytase and its complex with
 RT phytate.";
 RL Nat. Struct. Biol. 7:108-113(2000).
 CC -1- CATALYTIC ACTIVITY: A phosphate monoester + H(2)O = an alcohol +
 CC phosphate.
 CC -1- CATALYTIC ACTIVITY: Myo-inositol hexakisphosphate + H(2)O = 1D-
 CC myo-inositol 1,2,3,4,5-pentakisphosphate + phosphate.
 CC -1- SUBUNIT: Monomer.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- INDUCTION: In addition to cAMP-mediated control, this enzyme is
 CC induced when bacterial cultures reach stationary phase; its
 CC synthesis is triggered by phosphate starvation or a shift from
 CC aerobic to anaerobic conditions.
 CC -1- SIMILARITY: Belongs to the histidine acid phosphatase family.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs license
 CC -----
 CC EMBL: M58708: AAA72086.1; -; Genomic DNA.
 CC EMBL: U00096: AAC74065.1; -; Genomic DNA.
 CC EMBL: D90735: BAA35745.1; -; Genomic DNA.
 CC EMBL: X05471: CAA29031.1; -; Genomic DNA.
 CC EMBL: M58708: -; NOT ANNOTATED CDS; Genomic DNA.
 CC EMBL: S63811: AAB20286.1; -; Genomic DNA.
 CC PIR: B36733: B36733.
 CC PDB: 1DKL: X-ray; A/B=23-432.
 CC PDB: 1DKM: X-ray; A=23-432.
 CC PDB: 1DKN: X-ray; A=23-432.
 CC PDB: 1DKO: X-ray; A=23-432.
 CC PDB: 1DKP: X-ray; A=23-432.

DR PDB: 1DKO: X-ray; A=23-432.
 DR SWISS-2DPAGE: P07102; COLI.
 DR GenomeReviews: U00096_GR; b0380.
 DR ECHOBASE: EB0047; -.
 DR EcoGene: EG10049; appa.
 DR BioCyc: EcoCyc:APPA-MONOMER; -.
 DR InterPro: IPR000560; HisAc phsphtse.
 DR Pfam: PF00328; Acid_phosphat_A; 1.
 DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; 1.
 DR 3D-structure; Complete proteome; Direct protein sequencing; Hydrolase;
 DR Multifunctional enzyme; Periplasmic; Signal.
 KW SIGNAL
 FT 1 22
 FT CHAIN 23 432
 FT
 FT ACT_SITE 39 39
 FT ACT_SITE 326 326
 FT DISULFID 99 130
 FT DISULFID 155 430
 FT DISULFID 200 210
 FT DISULFID 404 413
 FT CONFLICT 51 66
 FT
 FT CONFLICT 75 76
 FT CONFLICT 112 112
 FT STRAND 27 38
 FT STRAND 41 42
 FT STRAND 45 45
 FT HELIX 49 53
 FT TURN 54 54
 FT STRAND 56 57
 FT STRAND 63 64
 FT TURN 66 67
 FT STRAND 69 69
 FT HELIX 71 90
 FT TURN 91 92
 FT STRAND 93 94
 FT STRAND 96 98
 FT TURN 102 104
 FT STRAND 105 109
 FT STRAND 111 112
 FT HELIX 113 126
 FT STRAND 127 127
 FT TURN 128 129
 FT STRAND 134 135
 FT STRAND 138 138
 FT TURN 140 141
 FT STRAND 142 142
 FT HELIX 145 147
 FT TURN 149 153
 FT STRAND 154 155
 FT HELIX 159 169
 FT TURN 170 172
 FT HELIX 174 179
 FT TURN 180 181
 FT HELIX 182 192
 FT TURN 193 193
 FT HELIX 194 196
 FT STRAND 197 197
 FT HELIX 198 201
 FT TURN 202 202
 FT STRAND 204 204
 FT TURN 206 207
 FT STRAND 208 208
 FT HELIX 212 215
 FT STRAND 216 216
 FT STRAND 220 222
 FT STRAND 224 225
 FT STRAND 227 229
 FT TURN 231 231
 FT HELIX 232 249
 FT TURN 250 250
 FT STRAND 252 253
 MODVTPEMPYKL -> NAGCHPRRMANLAKT (in
 Ref. 3).
 EL -> DV (in Ref. 4).
 D -> S (in Ref. 4).
 Periplasmic appa protein.
 /Frid=PRO.0000023947.
 Nucleophile.
 Proton donor.

FT HELIX 254 257
FT TURN 258 258

Query Match 96.6%; Score 2182; DB 1; Length 432;
Best Local Similarity 98.1%; Pred. No. 9.3e-162;
Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

QY 1 MKAIIIPFLSLIPLTPQSAFQSBELKESVIVSRGVAPPAKATQMDVDPDAMP 60
DB 1 MKAIIIPFLSLIPLTPQSAFQSBELKESVIVSRGVAPPAKATQMDVDPDAMP 60
QY 61 TWPVKGELTPRGSELIAVIGHYMRQRLVADGLPKCGCGPSGOVAIIADVDETRKTGE 120
DB 61 TWPVKGELTPRGSELIAVIGHYMRQRLVADGLPKCGCGPSGOVAIIADVDETRKTGE 120
QY 121 APAAGIAPDCAITVHTQADTSSPDPLFNPDKTGVCOLDNANVTDAIISRAGSIADFTGH 180
DB 121 APAAGIAPDCAITVHTQADTSSPDPLFNPDKTGVCOLDNANVTDAIISRAGSIADFTGH 180
QY 181 YGTAFRELERVLPQSNCLKREKQDECSLTQALPSBLKVSADCVSLTGAVSLASMT 240
DB 181 YGTAFRELERVLPQSNCLKREKQDECSLTQALPSBLKVSADCVSLTGAVSLASMT 240
QY 241 EIFFLQQAQGMPEPGMGRITDISHQNTLLSLHNAQFDLQRTPEVARSATPLDLIKTA 300
DB 241 EIFFLQQAQGMPEPGMGRITDISHQNTLLSLHNAQFDLQRTPEVARSATPLDLIKTA 300
QY 301 LTPHPPOKQAYGVTLPTSVLFIAGHDTNLANIGALELNMWTLPGQPDNTPPGSELVFERM 360
DB 301 LTPHPPOKQAYGVTLPTSVLFIAGHDTNLANIGALELNMWTLPGQPDNTPPGSELVFERM 360
QY 361 RRLSNSQMIQVSLVFQTLQOMRDKTPISLNTPPGEVKLTLAGCERNNAQMCISLAGFTQ 420
DB 361 RRLSNSQMIQVSLVFQTLQOMRDKTPISLNTPPGEVKLTLAGCERNNAQMCISLAGFTQ 420
QY 421 IYNEARIPACSL 432
DB 421 IYNEARIPACSL 432

```

RESULT 2
Q323E1 SHISS PRELIMINARY; PRT; 432 AA.
ID Q323E1 SHISS
AC Q323E1
DT 27-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 27-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 5.
DE Phosphoanhydride phosphorylase.
GN Name=appa; OrderedLocustNames=SSO_0987; ORFNames=SSO_0987;
OS Shigella sonnei (strain S9046).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=300269;
RN [1]
RP NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA).
RX PubMed-16275786; DOI=10.1093/nar/gki954;
RA Yang F., Yang J., Zhang X., Chen L., Jiang Y., Yan Y., Tang X.,
RA Wang J., Xiong Z., Dong J., Xue Y., Zhu Y., Xu X., Sun L., Chen S.,
RA Nie H., Peng J., Xu J., Wang Y., Yuan Z., Wen Y., Yao Z., Shen Y.,
RA Qiang B., Hou Y., Yu J., Jin Q.;
RT "Genome dynamics and diversity of Shigella species, the etiologic
RT agents of bacillary dysentery";
RL Nucleic Acids Res. 33:6445-6458(2005).
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC EMBL; CP0000038; AA287721.1; -; Genomic DNA.
CC GO; GO:0003993; F:acid phosphatase activity; IEA.
DR InterPro; IPR000560; HsAc_pheptase.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.

KW Complete proteome.
SQ SEQUENCE 432 AA; 47057 MW; 6510C6C579177F11 CRC64;

Query Match 96.6%; Score 2182; DB 2; Length 432;
Best Local Similarity 98.1%; Pred. No. 9.3e-162;
Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

QY 1 MKAIIIPFLSLIPLTPQSAFQSBELKESVIVSRGVAPPAKATQMDVDPDAMP 60
DB 1 MKAIIIPFLSLIPLTPQSAFQSBELKESVIVSRGVAPPAKATQMDVDPDAMP 60
QY 61 TWPVKGELTPRGSELIAVIGHYMRQRLVADGLPKCGCGPSGOVAIIADVDETRKTGE 120
DB 61 TWPVKGELTPRGSELIAVIGHYMRQRLVADGLPKCGCGPSGOVAIIADVDETRKTGE 120
QY 121 APAAGIAPDCAITVHTQADTSSPDPLFNPDKTGVCOLDNANVTDAIISRAGSIADFTGH 180
DB 121 APAAGIAPDCAITVHTQADTSSPDPLFNPDKTGVCOLDNANVTDAIISRAGSIADFTGH 180
QY 181 YGTAFRELERVLPQSNCLKREKQDECSLTQALPSBLKVSADCVSLTGAVSLASMT 240
DB 181 YGTAFRELERVLPQSNCLKREKQDECSLTQALPSBLKVSADCVSLTGAVSLASMT 240
QY 241 EIFFLQQAQGMPEPGMGRITDISHQNTLLSLHNAQFDLQRTPEVARSATPLDLIKTA 300
DB 241 EIFFLQQAQGMPEPGMGRITDISHQNTLLSLHNAQFDLQRTPEVARSATPLDLIKTA 300
QY 301 LTPHPPOKQAYGVTLPTSVLFIAGHDTNLANIGALELNMWTLPGQPDNTPPGSELVFERM 360
DB 301 LTPHPPOKQAYGVTLPTSVLFIAGHDTNLANIGALELNMWTLPGQPDNTPPGSELVFERM 360
QY 361 RRLSNSQMIQVSLVFQTLQOMRDKTPISLNTPPGEVKLTLAGCERNNAQMCISLAGFTQ 420
DB 361 RRLSNSQMIQVSLVFQTLQOMRDKTPISLNTPPGEVKLTLAGCERNNAQMCISLAGFTQ 420
QY 421 IYNEARIPACSL 432
DB 421 IYNEARIPACSL 432

```

RESULT 3
Q8GN88 ECOLI PRELIMINARY; PRT; 432 AA.
ID Q8GN88 ECOLI
AC Q8GN88;
DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2003, sequence version 1.
DT 07-FEB-2006, entry version 11.
DE Appa.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX Zhang Z., He J., Yao B., Zhou Y., Chen Y., Yi Y.;
RL Submitted (AUG-2002) to the EMBL/Genbank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX Chen Y., Zhu Z., Zhang Z., He J.;
RL Submitted (AUG-2002) to the EMBL/Genbank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC EMBL; AF537219; AAN28334.1; -; Genomic DNA.
CC HSSP; P07102; 1DKM.
DR SMR; Q8GN88; 23-432.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR InterPro; IPR000560; HsAc_pheptase.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.

SEQ SEQUENCE 432 AA; 47056 MW; 58355D76E7377737 CRC64;

Query Match 96.5%; Score 2178; DB 2; Length 432;
Best Local Similarity 97.9%; Pred. No. 1,9e-161;
Matches 423; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 MKAILPPLSLILPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
DB 1 MKAILPPLSLILPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60

QY 61 TWPVKLGELTPRGSELIAVIGHYWRQRLVADGLPKGCGPQSGQVAIIADVDETRKTGE 120
DB 61 TWPVKLGELTPRGSELIAVIGHYWRQRLVADGLPKGCGPQSGQVAIIADVDETRKTGE 120

QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTVGCQLDNNAVTDAILSRAGGSTADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTVGCQLDNNAVTDAILSRAGGSTADFTGH 180

QY 181 YQTAFFRELERYLNPPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
DB 181 RQTAFFRELERYLNPPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240

QY 241 EIFLLQAOAGMPBEQMGRIITDSHOMNTLLSLHNAQFDLLQRTPEVARSRAATPLDLIKTA 300
DB 241 EIFLLQAOAGMPBEQMGRIITDSHOMNTLLSLHNAQFDLLQRTPEVARSRAATPLDLIKTA 300

QY 301 LTPHPQKQAGVTLPTSVLPIAGHDNTLANLGALBLNMTLLPGQPDNTPPGSELVPERW 360
DB 301 LTPHPQKQAGVTLPTSVLPIAGHDNTLANLGALBLNMTLLPGQPDNTPPGSELVPERW 360

QY 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPSLNTPPGEVKTLTAGCEERNAQMGSLAGFTQ 420
DB 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPSLNTPPGEVKTLTAGCEERNAQMGSLAGFTQ 420

QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 4
Q8RKD6_ECOLI PRELIMINARY; PRT; 432 AA.
AC Q8RKD6;
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2002, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Periplasmic phosphoanhydride phosphohydrolase.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;

RM NUCLEOTIDE SEQUENCE.
RP MEDLINE=9036816; PubMed=2168385;
RA Dassa J., Marcq C., Boquet P.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals a significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase.";
RL J. Bacteriol. 172:5497-5500(1990).
RN [1]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=93054596; PubMed=1429631;
RA Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.-M.,
RA van Ertten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase.";
RL J. Biol. Chem. 267:22830-22836(1992).
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC EMBL; L03374; AAA00006.1; -; Genomic_DNA.

DR HSSP; P07102; 1DKM.
DR SMR; Q8RKD6; 23-432.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR00560; HisAc phosphatase.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolase.
SQ SEQUENCE 432 AA; 46972 MW; AF6C41E6193A5 CRC64;

Query Match 96.4%; Score 2176; DB 2; Length 432;
Best Local Similarity 97.9%; Pred. No. 2,7e-161;
Matches 423; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKAILPPLSLILPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
DB 1 MKAILPPLSLILPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60

QY 61 TWPVKLGELTPRGSELIAVIGHYWRQRLVADGLPKGCGPQSGQVAIIADVDETRKTGE 120
DB 61 TWPVKLGELTPRGSELIAVIGHYWRQRLVADGLPKGCGPQSGQVAIIADVDETRKTGE 120

QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTVGCQLDNNAVTDAILSRAGGSTADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTVGCQLDNNAVTDAILSRAGGSTADFTGH 180

QY 181 YQTAFFRELERYLNPPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
DB 181 RQTAFFRELERYLNPPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240

QY 241 EIFLLQAOAGMPBEQMGRIITDSHOMNTLLSLHNAQFDLLQRTPEVARSRAATPLDLIKTA 300
DB 241 EIFLLQAOAGMPBEQMGRIITDSHOMNTLLSLHNAQFDLLQRTPEVARSRAATPLDLIKTA 300

QY 301 LTPHPQKQAGVTLPTSVLPIAGHDNTLANLGALBLNMTLLPGQPDNTPPGSELVPERW 360
DB 301 LTPHPQKQAGVTLPTSVLPIAGHDNTLANLGALBLNMTLLPGQPDNTPPGSELVPERW 360

QY 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPSLNTPPGEVKTLTAGCEERNAQMGSLAGFTQ 420
DB 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPSLNTPPGEVKTLTAGCEERNAQMGSLAGFTQ 420

QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 5
Q8RKD7_ECOLI PRELIMINARY; PRT; 432 AA.
AC Q8RKD7;
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2002, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Periplasmic phosphoanhydride phosphohydrolase.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;

RM NUCLEOTIDE SEQUENCE.
RP MEDLINE=9036816; PubMed=2168385;
RA Dassa J., Marcq C., Boquet P.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals a significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase.";
RL J. Bacteriol. 172:5497-5500(1990).
RN [1]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=93054596; PubMed=1429631;
RA Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.-M.,

RA van Etten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase.";
J. Biol. Chem. 267:22830-22836(1992).
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
CC EMBL: L03373; AAA00005.1; -; Genomic_DNA.
DR HSSP: P07102; 1DKM.
DR SMR: QBRKD7; 23-432.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000560; H1aac phosphatase.
DR Pfam; PF00328; Acid phosphat A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
DR Hydrolyase.
KM
SQ SEQUENCE 432 AA; 46972 MW; 755D5E4B1AD916A6 CRC64;

Query Match 96.4%; Score 2176; DB 2; Length 432;
Best Local Similarity 97.9%; Pred. No. 2.7e-161;
Matches 423; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKAILIPFLSLIPLTPQSAFAQSEBELKESVIVSRHGVAAPTKATQLMQDVTDPDMP 60
DB 1 MKAILIPFLSLIPLTPQSAFAQSEBELKESVIVSRHGVAAPTKATQLMQDVTDPDMP 60
QY 61 TWPVKGLTPRGGLIAYLGHYRQRLVADGLPKCCGPOSGQVAIADVDERTKTGE 120
DB 61 TWPVKGLTPRGGLIAYLGHYRQRLVADGLPKCCGPOSGQVAIADVDERTKTGE 120
QY 121 AFAAGLAPDCAITYHTQADTSSPDLFNPPLKTVGQQLDNNAVTAIISRAGSIAIDFTGH 180
DB 121 AFAAGLAPDCAITYHTQADTSSPDLFNPPLKTVGQQLDNNAVTAIISRAGSIAIDFTGH 180
QY 181 YQTAFRELERLVLPQSNLCLKREKODSCSLTQALPSELKVSADCVSLTGAVALASMLT 240
DB 181 YQTAFRELERLVLPQSNLCLKREKODSCSLTQALPSELKVSADCVSLTGAVALASMLT 240
QY 241 EIFLLQQAQMGPEBGMGRITDSHQNNTLLSIHNAQFDLLQRTPEVARSRAATPLDLIKTA 300
DB 241 EIFLLQQAQMGPEBGMGRITDSHQNNTLLSIHNAQFDLLQRTPEVARSRAATPLDLIKTA 300
QY 301 LTPHPQQAQVGTLPSTVLFIAQHDNTLANLGALBELNMTLPQGPNDTPPGSELVFERW 360
DB 301 LTPHPQQAQVGTLPSTVLFIAQHDNTLANLGALBELNMTLPQGPNDTPPGSELVFERW 360
QY 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPSLNTPGSEVLTLAGCEERNAQMGCSLAGFTQ 420
DB 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPSLNTPGSEVLTLAGCEERNAQMGCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 6
QBRKD8_ECOLI PRELIMINARY; PRT; 432 AA.
AC QBRKD8;
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2002, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Periplasmic phosphoanhydride phosphohydrolase.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=30368616; PubMed=2168385;

RA Dasra J., Marck C., Boquet P.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase.";
J. Bacteriol. 172:5497-5500(1990).
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
CC EMBL: L03372; AAA00004.1; -; Genomic_DNA.
DR HSSP: P07102; 1DKM.
DR SMR: QBRKD8; 23-432.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000560; H1aac phosphatase.
DR Pfam; PF00328; Acid phosphat A; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
DR Hydrolyase.
KM
SQ SEQUENCE 432 AA; 46972 MW; 9A85536B57FCCFB5 CRC64;

Query Match 96.4%; Score 2176; DB 2; Length 432;
Best Local Similarity 97.9%; Pred. No. 2.7e-161;
Matches 423; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKAILIPFLSLIPLTPQSAFAQSEBELKESVIVSRHGVAAPTKATQLMQDVTDPDMP 60
DB 1 MKAILIPFLSLIPLTPQSAFAQSEBELKESVIVSRHGVAAPTKATQLMQDVTDPDMP 60
QY 61 TWPVKGLTPRGGLIAYLGHYRQRLVADGLPKCCGPOSGQVAIADVDERTKTGE 120
DB 61 TWPVKGLTPRGGLIAYLGHYRQRLVADGLPKCCGPOSGQVAIADVDERTKTGE 120
QY 121 AFAAGLAPDCAITYHTQADTSSPDLFNPPLKTVGQQLDNNAVTAIISRAGSIAIDFTGH 180
DB 121 AFAAGLAPDCAITYHTQADTSSPDLFNPPLKTVGQQLDNNAVTAIISRAGSIAIDFTGH 180
QY 181 YQTAFRELERLVLPQSNLCLKREKODSCSLTQALPSELKVSADCVSLTGAVALASMLT 240
DB 181 YQTAFRELERLVLPQSNLCLKREKODSCSLTQALPSELKVSADCVSLTGAVALASMLT 240
QY 241 EIFLLQQAQMGPEBGMGRITDSHQNNTLLSIHNAQFDLLQRTPEVARSRAATPLDLIKTA 300
DB 241 EIFLLQQAQMGPEBGMGRITDSHQNNTLLSIHNAQFDLLQRTPEVARSRAATPLDLIKTA 300
QY 301 LTPHPQQAQVGTLPSTVLFIAQHDNTLANLGALBELNMTLPQGPNDTPPGSELVFERW 360
DB 301 LTPHPQQAQVGTLPSTVLFIAQHDNTLANLGALBELNMTLPQGPNDTPPGSELVFERW 360
QY 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPSLNTPGSEVLTLAGCEERNAQMGCSLAGFTQ 420
DB 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPSLNTPGSEVLTLAGCEERNAQMGCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 7
QBRKD8_ECOLI PRELIMINARY; PRT; 432 AA.
AC QBRKD8;
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2002, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Periplasmic phosphoanhydride phosphohydrolase.

```

GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
[1]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=9036816; PubMed=2168385;
RA Dassa J., Marck C., Boquet P.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase."
RL J. Bacteriol. 172:5497-5500 (1990).
[2]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=93054596; PubMed=1429631;
RA Ostranin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.-M.,
RA van Eeten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase."
RL J. Biol. Chem. 267:22830-22836 (1992).
-----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivative License
-----
DR EMBL; I03370; AAA00002.1; -; Genomic_DNA.
DR HSSP; P07102; IDKQ.
DR SMK; Q8RKD9; 23-432.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR InterPro; IPR00560; HisAc_phsphtse.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KM Hydrolyase.
SQ SEQUENCE 432 AA; 46972 MW; 5BBD632D4682EABF CRC64;

Query Match 96.4%; Score 2176; DB 2; Length 432;
Best Local Similarity 97.9%; Pred. No. 2.7e-161;
Matches 423; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKAILPFLSLILPLTPQSAFAQSEPELKESVVIIVSRHGVRAPTKATQTMQDVTDPAMP 60
DB 1 MKAILPFLSLILPLTPQSAFAQSEPELKESVVIIVSRHGVRAPTKATQTMQDVTDPAMP 60

QY 61 TWPVKLGELTPRGSELAIYIGHYWRORLVADGLPKGCGQSGQVAIIADVDETRTKGE 120
DB 61 TWPVKLGELTPRGSELAIYIGHYWRORLVADGLPKGCGQSGQVAIIADVDETRTKGE 120

QY 121 AFAAGLAPDCAITHTQADTSSPDPLFNPLKTGVCOQDNNAVTDAILERAGSIADFTGH 180
DB 121 AFAAGLAPDCAITHTQADTSSPDPLFNPLKTGVCOQDNNAVTDAILERAGSIADFTGH 180

QY 181 YQTAFFRELERYLNFPQSNCLIKREKODESCSLTQALPSELKVSADCVSLTGAVALSMLT 240
DB 181 YQTAFFRELERYLNFPQSNCLIKREKODESCSLTQALPSELKVSADCVSLTGAVALSMLT 240

QY 241 EIFLLQQAQGMPEPGWGRITDSSHQNTLLSLHNAQFDLLQRTPEVASRATPLDLDKTA 300
DB 241 EIFLLQQAQGMPEPGWGRITDSSHQNTLLSLHNAQFDLLQRTPEVASRATPLDLDKTA 300

QY 301 LTPHPPOKAYGVTLPTSVLFIAGHDTNLANIAGALBLNMTLPQGPNDTPPGSELVFERW 360
DB 301 LTPHPPOKAYGVTLPTSVLFIAGHDTNLANIAGALBLNMTLPQGPNDTPPGSELVFERW 360

QY 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPGGEVLTLAGCEERNAQMCISLAGFTQ 420
DB 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPGGEVLTLAGCEERNAQMCISLAGFTQ 420

QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

```

```

RESULT 8
Q8RKD9_ECOLI PRELIMINARY; PRT; 432 AA.
ID Q8RKD9_ECOLI
AC Q8RKD9;
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2002, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Periplasmic phosphoanhydride phosphohydrolyase.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
[1]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=9036816; PubMed=2168385;
RA Dassa J., Marck C., Boquet P.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase."
RL J. Bacteriol. 172:5497-5500 (1990).
[2]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=93054596; PubMed=1429631;
RA Ostranin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.-M.,
RA van Eeten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase."
RL J. Biol. Chem. 267:22830-22836 (1992).
-----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivative License
-----
DR EMBL; I03371; AAA00003.1; -; Genomic_DNA.
DR HSSP; P07102; IDKQ.
DR SMK; Q8RKD9; 23-432.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR InterPro; IPR00560; HisAc_phsphtse.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KM Hydrolyase.
SQ SEQUENCE 432 AA; 47034 MW; 9F29B9DF9C368175 CRC64;

Query Match 96.3%; Score 2175; DB 2; Length 432;
Best Local Similarity 97.9%; Pred. No. 3.3e-161;
Matches 423; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 MKAILPFLSLILPLTPQSAFAQSEPELKESVVIIVSRHGVRAPTKATQTMQDVTDPAMP 60
DB 1 MKAILPFLSLILPLTPQSAFAQSEPELKESVVIIVSRHGVRAPTKATQTMQDVTDPAMP 60

QY 61 TWPVKLGELTPRGSELAIYIGHYWRORLVADGLPKGCGQSGQVAIIADVDETRTKGE 120
DB 61 TWPVKLGELTPRGSELAIYIGHYWRORLVADGLPKGCGQSGQVAIIADVDETRTKGE 120

QY 121 AFAAGLAPDCAITHTQADTSSPDPLFNPLKTGVCOQDNNAVTDAILERAGSIADFTGH 180
DB 121 AFAAGLAPDCAITHTQADTSSPDPLFNPLKTGVCOQDNNAVTDAILERAGSIADFTGH 180

QY 181 YQTAFFRELERYLNFPQSNCLIKREKODESCSLTQALPSELKVSADCVSLTGAVALSMLT 240
DB 181 YQTAFFRELERYLNFPQSNCLIKREKODESCSLTQALPSELKVSADCVSLTGAVALSMLT 240

QY 241 EIFLLQQAQGMPEPGWGRITDSSHQNTLLSLHNAQFDLLQRTPEVASRATPLDLDKTA 300
DB 241 EIFLLQQAQGMPEPGWGRITDSSHQNTLLSLHNAQFDLLQRTPEVASRATPLDLDKTA 300

QY 301 LTPHPPOKAYGVTLPTSVLFIAGHDTNLANIAGALBLNMTLPQGPNDTPPGSELVFERW 360
DB 301 LTPHPPOKAYGVTLPTSVLFIAGHDTNLANIAGALBLNMTLPQGPNDTPPGSELVFERW 360

QY 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPGGEVLTLAGCEERNAQMCISLAGFTQ 420
DB 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPGGEVLTLAGCEERNAQMCISLAGFTQ 420

```

```

Db      361 RRLSDNSQMIQVSLVFQTLQWRDXTPLSLNTPPEGVKLTLAGCERNNAQGCISLAGFTQ 420
Qy      421 IVNEARIPACSL 432
Db      421 IVNEARIPACSL 432

RESULT 9
Q8KRD5_ECOLI PRELIMINARY; PRT; 432 AA.
AC Q8KRD5;
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2002, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE periplasmic phosphoanhydride phosphohydrolase.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxId=562;

NUCLEOTIDE SEQUENCE.
RP MEDLINE=90368616; PubMed=2168385;
RA Dassa J., Marck C., Boquet P.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appa
RT glucose-1-phosphatase."
RL J. Bacteriol. 172:5497-5500 (1990).
[2]

NUCLEOTIDE SEQUENCE.
RP MEDLINE=93054596; PubMed=1429631;
RA Ostalan K., Harms E.H., Stevia P.B., Kuciel R., Zhou M.-M.,
RA van Eten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase."
RL J. Biol. Chem. 267:22830-22836 (1992).

-----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
CC EMBL: L03375; AAA00007.1; -; Genomic DNA.
CC
DR HSSP; P07102; 1DKM.
DR SMR; Q8KRD5; 23-432.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000560; HsPAC.phsphtse.
DR Pfam; PF00328; Acid.phosphat.A; 1.
DR PROSITE; PS00616; HTS_ACID_PHOSPHAT_1; 1.
KM Hydrolyase.
SQ SEQUENCE 432 AA; 46991 MW; 951F393EA9A1A47C CRC64;

Query Match 96.2%; Score 2172; DB 2; Length 432;
Best Local Similarity 97.9%; Pred. No. 5,6e-161;
Matches 423; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy      1 MKAILIPLSLILPLTPQSAFAQSEPELKESVIVSHGVRAPTKATQLMQDVTTPDAMP 60
Db      1 MKAILIPLSLILPLTPQSAFAQSEPELKESVIVSHGVRAPTKATQLMQDVTTPDAMP 60
Qy      61 TWPKLGLTPRGSELIAVLYGHYWRQRLVADGLLPKCCGPOSGOVAIADVDERTRKTGE 120
Db      61 TWPKLGLTPRGSELIAVLYGHYWRQRLVADGLLPKCCGPOSGOVAIADVDERTRKTGE 120
Qy      121 AFAAGLAPDCAIYHTQADTSSPDLFNPPLKTVGVCQLDNANVTDAIIRAGGSIAIDFTGH 180
Db      121 AFAAGLAPDCAIYHTQADTSSPDLFNPPLKTVGVCQLDNANVTDAIIRAGGSIAIDFTGH 180
Qy      181 YQTAFRELERLVANPQSNLCLKREKQDESGSLTQALPSELKVSADCVSLTGAVSLASMLT 240
Db      181 YQTAFRELERLVANPQSNLCLKREKQDESGSLTQALPSELKVSADCVSLTGAVSLASMLT 240
Qy      241 EIFLLQQAQGMPEBQWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRAATPLDLIKTA 300

```

```

Db      241 EIFLLQQAQGMPEBQWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRAATPLDLIKTA 300
Qy      301 LTPHPQKQAVGVTLPTSVLFIAGHDTNLAMGLALINMTLPQDPNTPPGSELVFERM 360
Db      301 LTPHPQKQAVGVTLPTSVLFIAGHDTNLAMGLALINMTLPQDPNTPPGSELVFERM 360
Qy      361 RRLSDNSQMIQVSLVFQTLQWRDXTPLSLNTPPEGVKLTLAGCERNNAQGCISLAGFTQ 420
Db      361 RRLSDNSQMIQVSLVFQTLQWRDXTPLSLNTPPEGVKLTLAGCERNNAQGCISLAGFTQ 420
Qy      421 IVNEARIPACSL 432
Db      421 IVNEARIPACSL 432

RESULT 10
Q31YP3_SHIBS PRELIMINARY; PRT; 432 AA.
AC Q31YP3;
DT 06-DEC-2005, integrated into UniProtKB/TrEMBL.
DT 06-DEC-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Phosphoanhydride phosphorylase.
GN Name=appa; OrderedLocustNames=SBO_2250;
OS Shigella boydii serotype 4 (strain Sb227).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxId=300268;
[1]

NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP PubMed=16275786; DOI=10.1093/nar/gk1954;
RA Yang F., Yang J., Zhang X., Chen L., Jiang Y., Yan Y., Tang X.,
RA Wang J., Xiong Z., Dong Z., Xue Y., Zhu Y., Xu X., Sun L., Chen S.,
RA Nie H., Peng J., Xu J., Wang Y., Yuan Z., Wen Y., Yao Z., Shen Y.,
RA Qiang B., Hou Y., Yu J., Jin Q.;
RT "Genome dynamics and diversity of Shigella species, the etiologic
RT agents of bacillary dysentery."
RL Nucleic Acids Res. 33:6445-6458 (2005).

-----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
CC EMBL: CP000036; AB86815.1; -; Genomic DNA.
CC
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
KM Complete proteome.
SQ SEQUENCE 432 AA; 47064 MW; 8569022BA2809C73 CRC64;

Query Match 95.4%; Score 2154; DB 2; Length 432;
Best Local Similarity 97.0%; Pred. No. 1.4e-159;
Matches 419; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy      1 MKAILIPLSLILPLTPQSAFAQSEPELKESVIVSHGVRAPTKATQLMQDVTTPDAMP 60
Db      1 MKAILIPLSLILPLTPQSAFAQSEPELKESVIVSHGVRAPTKATQLMQDVTTPDAMP 60
Qy      61 TWPKLGLTPRGSELIAVLYGHYWRQRLVADGLLPKCCGPOSGOVAIADVDERTRKTGE 120
Db      61 TWPKLGLTPRGSELIAVLYGHYWRQRLVADGLLPKCCGPOSGOVAIADVDERTRKTGE 120
Qy      121 AFAAGLAPDCAIYHTQADTSSPDLFNPPLKTVGVCQLDNANVTDAIIRAGGSIAIDFTGH 180
Db      121 AFAAGLAPDCAIYHTQADTSSPDLFNPPLKTVGVCQLDNANVTDAIIRAGGSIAIDFTGH 180
Qy      181 YQTAFRELERLVANPQSNLCLKREKQDESGSLTQALPSELKVSADCVSLTGAVSLASMLT 240
Db      181 YQTAFRELERLVANPQSNLCLKREKQDESGSLTQALPSELKVSADCVSLTGAVSLASMLT 240
Qy      241 EIFLLQQAQGMPEBQWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRAATPLDLIKTA 300
Db      241 EIFLLQQAQGMPEBQWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRAATPLDLIKTA 300
Qy      301 LTPHPQKQAVGVTLPTSVLFIAGHDTNLAMGLALINMTLPQDPNTPPGSELVFERM 360

```

```

Db      ||| 301 LTPHPQKQAYGVTLPFSVFIAGHDNTNLNLGALFLNMTLPQDPNTPPGGSLVFERW 360
Qy      ||| 361 RRLSDNSQWIOVSLVFQTLQOMDKTPLSLNTPPGGEVYKTLTACGEERNAQMGSLAGFTQ 420
Db      ||| 361 RRLSDNSQWIOVSLVFQTLQOMDKTPLSLNTPPGGEVYKTLTACGEERNAQMGSLAGFTQ 420
Qy      ||| 421 IYNEARIPACSL 432
Db      ||| 421 IYNEARIPACSL 432

RESULT 11
ID 063RW2_SHIFL PRELIMINARY; PRT; 432 AA.
AC 063RW2_
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2005, sequence version 4.
DE Phosphoanhydride phosphorylase; PH 2.5 acid phosphatase.
GN Name=appa; OrderedLocustNames=SF0982;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC SRRPAIN=301 / Serotype 2a; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang Y., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
CC EMBL; AE005674; AAN42610.2; -; Genomic_DNA.
CC SMR; Q83RW2; 23-432.
CC DR BIOCYC; SFLB198214:ANA42610.1-MONOMER; -.
CC DR GO; GO:0003993; F:acid phosphatase activity; IEA.
CC DR InterPro; IPR000560; HisAc_phosphatase.
CC DR Pfam; PF00328; Acid_phosphatase_2.
CC DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN_1.
CC DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
CC KM Complete proteome.
SQ SEQUENCE 432 AA; 47106 MW; 25093A5392B9C18C CRC64;

Query Match 95.2%; Score 2149; DB 2; Length 432;
Best Local Similarity 96.8%; Pred. No. 3.5e-159;
Matches 418; Conservative 1; Mismatches 13; Indels 0; Gaps 0;

```

```

Db      ||| 241 EIFLQQAQAMPBEPGWRITDISHQWNTLLSLHNAQFLLQRTPEVARSATPLLDLIMAA 300
Qy      ||| 301 LTPHPQKQAYGVTLPFSVFIAGHDNTNLNLGALFLNMTLPQDPNTPPGGSLVFERW 360
Db      ||| 301 LTPHPQKQAYGVTLPFSVFIAGHDNTNLNLGALFLNMTLPQDPNTPPGGSLVFERW 360
Qy      ||| 361 RRLSDNSQWIOVSLVFQTLQOMDKTPLSLNTPPGGEVYKTLTACGEERNAQMGSLAGFTQ 420
Db      ||| 361 RRLSDNSQWIOVSLVFQTLQOMDKTPLSLNTPPGGEVYKTLTACGEERNAQMGSLAGFTQ 420
Qy      ||| 421 IYNEARIPACSL 432
Db      ||| 421 IYNEARIPACSL 432

RESULT 12
ID 06RK08_ECOLI PRELIMINARY; PRT; 432 AA.
AC 06RK08_
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DE 07-FEB-2006, entry version 8.
DE Acid phosphatase/phytase 2.
GN Name=appa2;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC MEDLINE=59194564; PubMed=10092520; DOI=10.1006/dbrc.1999.0361;
RA Rodriguez E., Han Y., Lei X.G.;
RT "Cloning, sequencing, and expression of an Escherichia coli acid
RT phosphatase/phytase gene (appa2) isolated from pig colon."
RL Biochem. Biophys. Res. Commun. 257:117-123(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Lei X.G.;
RL Submitted (DEC-2003) to the EMBL/Genbank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
CC EMBL; AY496073; AAR87658.1; -; Genomic_DNA.
CC SMR; 06RK08; 23-432.
CC DR GO; GO:0003993; F:acid phosphatase activity; IEA.
CC DR InterPro; IPR000560; HisAc_phosphatase.
CC DR Pfam; PF00328; Acid_phosphatase_1.
CC DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
CC DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
CC FT CHAIN 20 432 acid phosphatase/phytase 2.
SQ SEQUENCE 432 AA; 47042 MW; 71B05EBE2EA2674 CRC64;

Query Match 95.1%; Score 2147; DB 2; Length 432;
Best Local Similarity 96.8%; Pred. No. 5.1e-159;
Matches 418; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

```

QY 241 EFLAQAGMPBPGRITTDHSHQNTLLSHNAQFDLQRTPEVARSPATPLDLIKTA 300
DB 241 EFLAQAGMPBPGRITTDHSHQNTLLSHNAQFDLQRTPEVARSPATPLDLIKTA 300
QY 301 LTPHPPOKQAYGVTLPTSVLFIAGHDTNLANIGALBELNMTLPGQPDNTPPGSELVFERW 360
DB 301 LTPHPPOKQAYGVTLPTSVLFIAGHDTNLANIGALBELNMTLPGQPDNTPPGSELVFERW 360
QY 361 RRLSDNSQMIQVSLVFQTLQQRDKTPISLNTPPGSEVKTLTLAGCERNNAQGCISLAGFTQ 420
DB 361 RRLSDNSQMIQVSLVFQTLQQRDKTPISLNTPPGSEVKTLTLAGCERNNAQGCISLAGFTQ 420
QY 421 IYNEARIPACSL 432
DB 421 IYNEARIPACSL 432

RESULT 13
ID Q7UD08_SHIFL PRELIMINARY; PRT; 432 AA.
AC Q7UD08;
DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2003, sequence version 1.
DT 07-MAR-2006, entry version 12.
DE Phosphoanhydride phosphorylase, pH 2.5 acid phosphatase.
GN Name=appa; OrderedLocustNames=S1048; ORFNames=S_1048;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA DOI=10.1128/JAI.71.5.2775-2786.2003;
RA Fountlerberg M.B., Burland V., Venkatesan M.M., Deng W.,
Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DB EMBL: AE014073; AAP16495.1; -; Genomic_DNA.
DR HSRP; P07102, 1DKL.
DR SMR; Q7UD08; 23-432.
DR GenomeReviews; AE014073 GR; S1048.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR InterPro; IPR000560; H18AC.phosphatase.
DR Pfam; PF00328; Acid.phosphat.A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN 1.
SQ SEQUENCE 432 AA; 47136 MW; 20044239BC6D6FC CRC64;

Query Match 95.0%; Score 2145; DB 2; Length 432;
Best Local Similarity 96.5%; Pred. No. 7.3e-159;
Matches 417; Conservative 1; Mismatches 14; Indels 0; Gaps 0;

QY 1 MKAILIPLSLILITPQSAFAQSEPELKLBSVIVSRHGVAPPTKATQLMDVTPDAMP 60
DB 1 MKAILIPLSLILITPQSAFAQSEPELKLBSVIVSRHGVAPPTKATQLMDVTPDAMP 60
QY 61 TWPVVKGLTTPRGSELIAVLYGHYWRQRLVADGLLPKCCGPGSGQVAIIVADVETRTKT 120
DB 61 TWPVVKGLTTPRGSELIAVLYGHYWRQRLVADGLLPKCCGPGSGQVAIIVADVETRTKT 120
QY 121 AFAAGIAPDCAITVTQADTSSPDLFNPPLKTVGQQLDNANVTDAIILBRAGSINDFTGH 180
DB 121 AFAAGIAPDCAITVTQADTSSPDLFNPPLKTVGQQLDNANVTDAIILBRAGSINDFTGH 180

QY 181 YQTAFLRELRLVINFQSNILCKREKDESSCSITQALPELKVASADCVSLTGAVSIASMLT 240
DB 181 YQTAFLRELRLVINFQSNILCKREKDESSCSITQALPELKVASADCVSLTGAVSIASMLT 240
QY 241 EFLAQAGMPBPGRITTDHSHQNTLLSHNAQFDLQRTPEVARSPATPLDLIKTA 300
DB 241 EFLAQAGMPBPGRITTDHSHQNTLLSHNAQFDLQRTPEVARSPATPLDLIKTA 300
QY 301 LTPHPPOKQAYGVTLPTSVLFIAGHDTNLANIGALBELNMTLPGQPDNTPPGSELVFERW 360
DB 301 LTPHPPOKQAYGVTLPTSVLFIAGHDTNLANIGALBELNMTLPGQPDNTPPGSELVFERW 360
QY 361 RRLSDNSQMIQVSLVFQTLQQRDKTPISLNTPPGSEVKTLTLAGCERNNAQGCISLAGFTQ 420
DB 361 RRLSDNSQMIQVSLVFQTLQQRDKTPISLNTPPGSEVKTLTLAGCERNNAQGCISLAGFTQ 420
QY 421 IYNEARIPACSL 432
DB 421 IYNEARIPACSL 432

RESULT 14
ID Q7AFW3_ECO57 PRELIMINARY; PRT; 434 AA.
AC Q7AFW3;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 10.
DE Phosphoanhydride phosphorylase.
GN OrderedLocustNames=BCa1136;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHRC;
RX MEDLINE=21156231; PubMed=11258796; DOI=10.1093/dnares/8.1.11;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tohe T.,
Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DB EMBL: BA000007; BAB34559.1; -; Genomic_DNA.
DR SMR; Q7AFW3; 25-434.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR InterPro; IPR000560; H18AC.phosphatase.
DR Pfam; PF00328; Acid.phosphat.A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN 1.
SQ SEQUENCE 434 AA; 47337 MW; F197DF7D1869F9C CRC64;

Query Match 95.0%; Score 2144; DB 2; Length 434;
Best Local Similarity 96.5%; Pred. No. 8.7e-159;
Matches 419; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

QY 1 MKAILIPLSLILITPQSAFAQSEPELKLBSVIVSRHGVAPPTKATQLMDVTPDAMP 58
DB 1 MKAILIPLSLILITPQSAFAQSEPELKLBSVIVSRHGVAPPTKATQLMDVTPDAMP 60
QY 59 WPTWPKVGLTTPRGSELIAVLYGHYWRQRLVADGLLPKCCGPGSGQVAIIVADVETRTKT 118
DB 61 WPTWPKVGLTTPRGSELIAVLYGHYWRQRLVADGLLPKCCGPGSGQVAIIVADVETRTKT 120
QY 119 GEARAGIAPDCAITVTQADTSSPDLFNPPLKTVGQQLDNANVTDAIILBRAGSINDFT 178
DB 119 GEARAGIAPDCAITVTQADTSSPDLFNPPLKTVGQQLDNANVTDAIILBRAGSINDFT 178

Db 121 GEAPFAGLADPCAITVHTQADTSSPDLEFNPDKTGVCOJLDNANVTDAILSRAGGSIA DFT 180
Qy 179 GHVOTARELERVYNPQSNCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASM 238
Db 181 GHRQTARELERVYNPQSNCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASM 240
Qy 239 LTFEFLQQAQMGPEPBGRIITDSHOWNTLLSLHNAQFDLQRTPEVARSRATPLDLIK 298
Db 241 LTFEFLQQAQMGPEPBGRIITDSHOWNTLLSLHNAQFDLQRTPEVARSRATPLDLIM 300
Qy 299 TALTPHPQKQAYGVTLPTSVLFTAGHDNTLANIGALIELNMTLPQGPNDTPPGGELVFE 358
Db 301 IALTTPHPQKQAYGVTLPTSVLFTAGHDNTLANIGALIELNMTLPQGPNDTPPGGELVFE 360
Qy 359 RMRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCERRNAQGCSLAGF 418
Db 361 RMRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCERRNAQGCSLAGF 420
Qy 419 TQIVNEARIPACSL 432
Db 421 TQIVNEARIPACSL 434

RESULT 15
Q8XC29_ECO57 PRELIMINARY; PRT; 444 AA.
ID 08XC29_ECO57
AC 08XC29
DT 01-MAR-2002, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 2.
DT 07-FEB-2006, entry version 17.
DB Phosphoamylidride phosphorylase; PH 2.5 acid phosphatase; periplasmic.
GN Name=appa; Ordered locus names=zl397;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=1126551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck B.J., Davis N.W., Lim A., Dimalanta B.T., Potamoudis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533 (2001).
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
EMBL; AE005174; AAG55528.1; -; Genomic_DNA.
DR PIR; D85633; D85633.
DR PIR; H90770; H90770.
DR HSSP; P07102; 1DKL.
DR SMR; O8XC29; 35-444.
DR BioCyc; ECO183334-1:EC51136-MONOMER; -;
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR InterPro; IPR000560; HisAc_phosphatase.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00718; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 444 AA; 48388 MW; A9AA7B653AF247B CRC64;

Query Match 95.0%; Score 2144; DB 2; Length 444;
Best Local Similarity 96.5%; Pred. No. 9e-159;
Matches 419; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

Qy 1 MKAILIPLSLIPLTPQSAFAQ--EPRLKESVIVSRHGVRAPYATATQMODVTPDA 58
Db 11 MKAILIPLSLIPLTPQSAFAQSEPEBELKLESVIVSRHGVRAPYATATQMODVTPDA 70

Qy 59 WPTWPVGLGELTPRGGLIAYLGHWYRQRLVADGLPEKCCPQSGQVAIIADYDERTRKT 118
Db 71 WPNWPVGLGMLTPRGGLIAYLGHWYRQRLVADGLPTKCCPQSGQVAIIADYDERTRKT 130
Qy 119 GEAPFAGLADPCAITVHTQADTSSPDLEFNPDKTGVCOJLDNANVTDAILSRAGGSIA DFT 178
Db 131 GEAPFAGLADPCAITVHTQADTSSPDLEFNPDKTGVCOJLDNANVTDAILSRAGGSIA DFT 190
Qy 179 GHVOTARELERVYNPQSNCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASM 238
Db 191 GHRQTARELERVYNPQSNCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASM 250
Qy 239 LTFEFLQQAQMGPEPBGRIITDSHOWNTLLSLHNAQFDLQRTPEVARSRATPLDLIK 298
Db 251 LTFEFLQQAQMGPEPBGRIITDSHOWNTLLSLHNAQFDLQRTPEVARSRATPLDLIM 310
Qy 299 TALTPHPQKQAYGVTLPTSVLFTAGHDNTLANIGALIELNMTLPQGPNDTPPGGELVFE 358
Db 311 IALTTPHPQKQAYGVTLPTSVLFTAGHDNTLANIGALIELNMTLPQGPNDTPPGGELVFE 370
Qy 359 RMRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCERRNAQGCSLAGF 418
Db 371 RMRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCERRNAQGCSLAGF 430
Qy 419 TQIVNEARIPACSL 432
Db 431 TQIVNEARIPACSL 444

Search completed: June 13, 2006, 10:28:33
Job time : 190.248 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 14, 2006, 12:41:52 ; Search time 220.459 Seconds

(without alignments)
5499.798 Million cell updates/sec

Title: US-10-601-319-10

Perfect score: 2258
Sequence: 1 MKALIPPLSLIPITPQSA.....CSLAFQIVNARIPACSL 432

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlp
-O=/abs/ABSSWB_Spool/US10601319/runat_12062006_180055_27578/app_query.fasta_1
-DB=Issued_Patents_NA -QFMT=fastest -SUFFIX=rm1 -MINMATCH=0.1 -LOOCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdt
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-HOST=abs803p -USER=US10601319 @CGN_1_1_381 @runat_12062006_180055_27578
-NCPU=6 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents_NA:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/ina/1.COMB.seq:*
- 2: /EMC_Celerra_SIDS3/ptodata/2/ina/5.COMB.seq:*
- 3: /EMC_Celerra_SIDS3/ptodata/2/ina/6A.COMB.seq:*
- 4: /EMC_Celerra_SIDS3/ptodata/2/ina/6B.COMB.seq:*
- 5: /EMC_Celerra_SIDS3/ptodata/2/ina/7.COMB.seq:*
- 6: /EMC_Celerra_SIDS3/ptodata/2/ina/H.COMB.seq:*
- 7: /EMC_Celerra_SIDS3/ptodata/2/ina/PC/US.COMB.seq:*
- 8: /EMC_Celerra_SIDS3/ptodata/2/ina/PP.COMB.seq:*
- 9: /EMC_Celerra_SIDS3/ptodata/2/ina/RG.COMB.seq:*
- 10: /EMC_Celerra_SIDS3/ptodata/2/ina/backfillseq.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2182	96.6	1901	US-09-866-379-7	Sequence 7, Appli
2	2182	96.6	1901	US-09-866-379-9	Sequence 9, Appli
3	2181	96.6	1901	US-09-866-379-5	Sequence 5, Appli
4	2175	96.3	1901	US-09-866-379-6	Sequence 6, Appli
5	2171	96.1	1323	US-09-259-214-1	Sequence 1, Appli
6	2171	96.1	1323	US-09-318-528-1	Sequence 1, Appli
7	2171	96.1	1323	US-09-291-931-1	Sequence 1, Appli
8	2171	96.1	1323	US-09-580-515-1	Sequence 1, Appli

9	2171	96.1	1323	3	US-09-866-379-1	Sequence 1, Appli
10	2164	95.8	1486	3	US-09-715-477-2	Sequence 2, Appli
11	2147	95.1	1489	3	US-09-540-149A-9	Sequence 9, Appli
12	2147	95.1	1489	4	US-10-266-041A-9	Sequence 9, Appli
13	2144	95.0	1486	3	US-09-715-477-4	Sequence 4, Appli
14	2063.5	91.4	1272	2	US-08-910-798-1	Sequence 1, Appli
15	543	24.0	1266	3	US-09-489-039A-341	Sequence 341, App
16	381	16.9	1569	3	US-09-489-039A-5330	Sequence 6330, Ap
17	356	15.8	5975	2	US-08-920-812-23	Sequence 23, Appl
18	356	15.8	5975	2	US-08-920-827-23	Sequence 23, Appl
19	356	15.8	5975	2	US-08-921-177-23	Sequence 23, Appl
20	356	15.8	5975	2	US-08-362-577C-23	Sequence 23, Appl
21	356	15.8	5975	2	US-08-920-828-23	Sequence 23, Appl
22	126	5.6	1571	3	US-08-868-435-32	Sequence 32, Appl
23	126	5.6	1571	3	US-08-744-231-32	Sequence 32, Appl
24	126	5.6	1571	3	US-09-044-718-10	Sequence 10, Appl
25	126	5.6	1571	3	US-09-635-504-32	Sequence 32, Appl
26	126	5.6	1571	3	US-10-062-848-10	Sequence 10, Appl
27	124	5.5	1404	3	US-09-684-855-166	Sequence 166, App
28	124	5.5	1404	3	US-09-488-265B-32	Sequence 32, Appl
29	117	5.2	7404	3	US-09-902-540-3115	Sequence 3115, Ap
30	117	5.2	18809	3	US-09-902-540-3115	Sequence 1141, Ap
31	116	5.1	1642	3	US-08-993-359-27	Sequence 27, Appl
32	116	5.1	1642	3	US-09-482-558A-27	Sequence 1, Appli
33	115	5.1	13508	3	US-09-655-270A-1	Sequence 1, Appli
34	115	5.1	12523	3	US-09-651-941-1	Sequence 1, Appli
35	115	5.1	12523	3	US-09-955-597-1	Sequence 1, Appli
36	114	5.0	28958	2	US-08-258-261B-6	Sequence 6, Appli
37	114	5.0	28958	2	US-08-456-837-6	Sequence 6, Appli
38	114	5.0	28958	2	US-08-457-342-6	Sequence 6, Appli
39	114	5.0	28958	2	US-08-457-646A-6	Sequence 6, Appli
40	114	5.0	28958	2	US-08-458-076A-6	Sequence 6, Appli
41	114	5.0	28958	2	US-08-764-233A-4	Sequence 4, Appli
42	114	5.0	28958	2	US-08-457-335A-6	Sequence 6, Appli
43	114	5.0	28958	2	US-08-729-214-6	Sequence 6, Appli
44	114	5.0	28958	3	US-09-028-934-6	Sequence 6, Appli
45	114	5.0	49377	2	US-08-764-233A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-866-379-7
Sequence 7, Application US/09866379
Patent No. 685365
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: SHORT, Jay
APPLICANT: KRETZ, Keith
APPLICANT: GRAY, Kevin
APPLICANT: BARTON, Nelson
APPLICANT: GARRETT, James
TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASIS AND USES THEREOF
FILE REFERENCE: DIVERSI370-7
CURRENT APPLICATION NUMBER: US/09/866,379
CURRENT FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: US 09/580,515
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 09/318,528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: US 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 09/259,214
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: US 08/910,798
PRIOR FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7
LENGTH: 1901
TYPE: DNA
ORGANISM: Escherichia coli

Best Local Similarity: 98.1% Mismatches: 8
 Query Match: 96.6% Indels: 0
 DB: 3 Gaps: 0

US-10-601-319-10 (1-432) x US-09-866-379-9 (1-1901)

```

QY      1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
DB      188 ATGAAGAAGCATCTTAATCCATTTTATCTCTTCATATCCGTTAAACCCGCAATTCGCA 247

QY      21 PheAlaGlnSerGluProGluLeuLeuGlnSerValValIleValSerAlaGly 40
DB      248 TTCGCTCAGAGTGAAGCCGAGCTGAAGAGTGAGTGTGATGTGTGATGCTGATGCTGAT 307

QY      41 ValAlaGlaProThrIleValAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
DB      308 GTGGGTGCTCCCAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 367

QY      61 ThrTTPProValIleLeuGlyIleuLeuThrProArgGlyGlyIleuLeuIleAlaTyrlau 80
DB      368 ACCTGGCCGGTAAACTGGGTGGCTGACACCCCGAGGTGGTGAAGCTTAATCCGCTATCTC 427

QY      81 GlyHisTTPArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
DB      428 GGACATTTACCAACGACGACGCTGGTGAAGCGACGATTCGCGCAAAAGGCGTGGCCG 487

QY      101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGlnArgThrArgIleThrGlyIle 120
DB      488 CAGTCTGGTCAAGGTGGCATTAATGCTGATGTCGACGACGACGACGACGACGACGACGAC 547

QY      121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
DB      548 GCCTTCGCGCCGCGGCTGGCACTGACTGCAATACCTGACATACCCGACGACGATGAC 607

QY      141 SerSerProAspProLeuPheAsnProLeuGlyThrGlyValCysGlnLeuAspAsnAla 160
DB      608 TCCAGTCCCGATCCGTTATTATCTCTTAATAACGCGGCTTGGCAATGATTAACGCG 667

QY      161 AsnValThrAspAlaIleLeuGlnArgAlaGlyIleSerIleAlaAspPheThrGlyHis 180
DB      668 AACGAGACTGACGACATCTCTGACGAGGACGAGGATGCAATGCTGACATTCGCGGCA 727

QY      181 TyrGlnThrAlaPheArgGluLeuGlnArgValLeuAsnPheProGlnSerAsnLeuCys 200
DB      728 CGGCAAAAGCGCTTCGCGAATCGAAGCGGTCTTAATTTCCGCAATCAACTGTCG 787

QY      201 LeuLysArgGluLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlnLeu 220
DB      788 CTTAAACGTGAAGAACAGGACGAAAGCTGTTCAATACGACGACATTCGAACTC 847

QY      221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
DB      848 AAGGAGACCGCGCAATGTCATTAACCGGTGGGTGAAGCTGCAATGCTGACG 907

QY      241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGlnProGlyThrGlyValArgIleThr 260
DB      908 GAGATATTTCTCTCCACACACGACGAGGATGCGGAGCGGGGTGGGAAAGGATCC 967

QY      261 AspSerHisGlnTTPAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
DB      968 GATTCACACCAAGTGAACACTTGTGAAGTTGCAATACGGCAATTTTGTGTTACA 1027

QY      281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
DB      1028 CGCAGCGCAAGAGTGGCCGCGCGCGCACCCGTTATTGAATTTGATTCAGACAGCG 1087

QY      301 LeuThrProHisProProGlnLysGlnAlaTyrlGlyValThrLeuProThrSerValLeu 320
DB      1088 TTGAGCGCCCATCCACCGCAAAACAGGCGTATGATGACATTCACCTTCAGTGGCTG 1147

QY      321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTTP 340
DB      1148 TTTATCGCGGACAGCATACTTAATCTGGCAATCTCGGCGGCGGACCTGGAGCTCAACTGG 1207
  
```

```

QY      341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyIleuValPheGlnArgTTP 360
DB      1208 ACGTTTCCCGGTCCACCGGATTAACACGCGCAGAGTGGTGAACGTGTGTGAACGCTGG 1267

QY      361 ArgArgLeuSerAspAsnSerGlnTTPIleGlnValSerLeuValPheGlnThrLeuGln 380
DB      1268 CCGTCCGCTAAGCATTAACGACGATGATTCAGGTTTCGCTGCTTCGACACTTTCAG 1327

QY      381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValIleLysLeuThr 400
DB      1328 CAGATGCGGATTAACGCGCTGTCATTAATACGCGCCGCGAGAGGTGAACATGAC 1387

QY      401 LeuAlaGlyCysGlnGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
DB      1388 CTGGCAGAGTGTGAAGAGGAAATGCGCAGGCGCATGTGTTCGTTGGCAGGTTTACGCA 1447

QY      421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
DB      1448 ATCGTGAATGAACACGCAATACCGCGCTGCACTTGG 1483
  
```

RESULT 3

US-09-866-379-5
 ; Sequence 5, Application US/09866379
 ; Patent No. 6855365

GENERAL INFORMATION:

APPLICANT: DIVERSA CORPORATION
 APPLICANT: SHORT, Jay
 APPLICANT: KREYER, Keith
 APPLICANT: GRAY, Kevin
 APPLICANT: BARRON, Nelson
 APPLICANT: GARRETT, James
 APPLICANT: O'DONOGHUE, Eileen
 TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
 FILE REFERENCE: DIVER1370-7
 CURRENT APPLICATION NUMBER: US/09/866,379
 CURRENT FILING DATE: 2001-05-24
 PRIOR APPLICATION NUMBER: US 09/580,515
 PRIOR FILING DATE: 2000-05-25
 PRIOR APPLICATION NUMBER: US 09/318,528
 PRIOR FILING DATE: 1999-05-25
 PRIOR APPLICATION NUMBER: US 09/291,931
 PRIOR FILING DATE: 1999-04-13
 PRIOR APPLICATION NUMBER: US 09/259,214
 PRIOR FILING DATE: 1999-03-01
 PRIOR APPLICATION NUMBER: US 08/910,798
 PRIOR FILING DATE: 1997-08-13
 NUMBER OF SEQ ID NOS: 10
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 5
 LENGTH: 1901
 TYPE: DNA
 ORGANISM: Escherichia coli
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (1)..(1901)
 OTHER INFORMATION: n is any nucleotide

US-09-866-379-5

Alignment Scores:

Pred. No.: 6.75e-223 Length: 1901
 Score: 2181.00 Matches: 424
 Percent Similarity: 98.1% Conservative: 0
 Best Local Similarity: 98.1% Mismatches: 8
 Query Match: 96.6% Indels: 0
 DB: 3 Gaps: 0

US-10-601-319-10 (1-432) x US-09-866-379-5 (1-1901)

```

QY      1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
DB      188 ATGAAGAAGCATCTTAATCCATTTTATCTCTTCATATCCGTTAAACCCGCAATTCGCA 247
  
```

OY	21	theaIaGInserGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly	40
Db	248	TTCCGTCAAGAGTGAAGCCGGAAGCTGAAGCTGGAAAGTGGTATGTCAGTCCATAGGT	307
OY	41	ValAAGValProThrLysIleAlaThrGlnLeuMetGlnAspValThrProAspAlaTyrPro	60
Db	308	GTGGCTGCTCCAAACCAAGGCCAAGCACTGAATGCAAGATGTCAACCCCAAGCATGGCCA	367
OY	61	ThrTyrProValLysLeuGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu	80
Db	368	ACCTGAGCGGTAAACCTGGGTGACTGACACACCGCGGTGGTGAAGCTAATCGCTATCTC	427
OY	81	GlyHisTyrTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro	100
Db	428	GGACATTCACCAACGGCAGCGGTGTGGTAAGCCAGATTCCTGGCGAAAAAGGGCTCCCG	487
OY	101	GlnSerGlyGlnValAlaIleIleAlaAspValAspGluAlaThrArgTyrThrGlyGlu	120
Db	488	CAGTCTGGTCAAGTGGCAATTAATGCTAATGTGCAGAGCGTAACCCGTAAACAGCGGAA	547
OY	121	AlaPheAlaIaIaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr	140
Db	548	GGCTTCGGCGCGGGGTGGCACTGACTGTGAATACCGTACATACCAAGGAGATAG	607
OY	141	SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla	160
Db	608	TCCAGTCCCGAATCCGTTATTTATCTCTAAAAACGGGGTTGGCCACCTGAGTAACGGG	667
OY	161	AsnValThrAspAlaIleLeuGlnLysArgAlaGlySerIleAlaAspPheThrGlyHis	180
Db	668	AACGTGACTGAAGCGCATCTTCAGCAGGCGAAGAGGTCATATGCTACCTTACCGGAGAT	727
OY	181	TyrGlnThrAlaPheArgGluLeuGlnArgValLeuAsnPheProGlnSerAsnLeuCys	200
Db	728	CGGCAAAACGGCGTTTCGGCAACTGGAAACGGGGTCTTAATTTTCCGCAATCAACTGGTG	787
OY	201	LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu	220
Db	788	CTTAAACGTGAAGAAACAGACCAAAAGCTGTTCATTAACACGAGCATTTACATCGGAATC	847
OY	221	LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerLeuThr	240
Db	848	AAGGTGAAGGCCGACATGTCTCATTTAACCGGTGGGTAAAGCTCCACATCAATGCTGAG	907
OY	241	GluIlePheLeuLeuGlnGlnAlaGlnIleCysMetProGluProGlyTyrTrpLysArgIleThr	260
Db	908	GAGATATTTCCTCTCGAACAACACACAGGAAATCCGCAACCGGGGGTGGGAAGATCAC	967
OY	261	AspSerHisGlnTyrAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln	280
Db	968	CATTCAACACAGGTGAACACCTTGGTCAATGTTGGATTAACGCGCAATTTATTTGCTAACAA	102
OY	281	ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla	300
Db	1028	CGCAGCCAGAGGTGGCCCGCAGCGCGCCACCCCGTATTAAGATTGATCAAGACAGG	108
OY	301	LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu	320
Db	1088	TTGAGCGCCCAATCCACCGCAAAACAGGGGTATGGTGAATACCATTCACAGTCAAGTGG	114
OY	321	PheIleIaIaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyValAlaLeuGluLeuAsnTyr	340
Db	1148	TTTATTCGCGGACAGATACTAATCTGGCAAAATCTCGCGGCGCACTGGAGCTCAACTGG	120
OY	341	ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluAlaGlyTyr	360
Db	1208	ACGCTTCCCGGTCAAGCCGGAATAACACGCCGCAAGGTGAAGCTGGTATTTGAACGCTGG	126
OY	361	ArgArgLeuSerAspAsnSerGlnTyrIleGlnValSerLeuValPheGlnThrLysGln	380
Db	1268	CGTCGGCTAAGAGTAACAGCACAGTGAATTCAGGTTTCGCTGGTCTTCACAGACTTTACAG	132
OY	381	GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr	400

```

Db      1338  CAGATGCTGATAAAACCCCGCTGCTATTAAATACGCCCGGAGAGGTGAATCTGACC 1387
Qy      401   Leu1aagiCySGluGluArGAsn1aGln1yMetCySerLeu1aGlyPheThrGln 420
Db      1388  CTGGCAGGATGTGGAAGAGCAAAATGCCGAGGGCATGTCTTGTTGGGACGATTTTAACGAA 1447
Qy      421   l1eValaanglu1aArG1lePro1aCySerleu 432
Db      1448  ATCGTGAATGAAGCAGCATACCGCGCTGCAGATTGG 1483

RESULT 4
US-09-866-379-6
; Sequence 6, Application US/09866379
; Patent No. 6855365
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Bileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-6

Alignment Scores:
Pred. No.:      2.96e-222      Length:      1901
Score:          2175.00      Matches:      423
Percent Similarity: 97.9%      Conservative: 0
Best Local Similarity: 97.9%      Mismatches: 9
Query Match:    96.3%      Indels:      0
DB:              3          Gaps:      0

US-10-601-319-10 (1-432) x US-09-866-379-6 (1-1901)
Qy      1      Met1ySa1a1le1e1le1Pro1he1uSer1e1u1le1Pro1e1uThr1Pro1Gln1Ser1a 20
Db      188  ATGAAAGGAACTTAATCCATTTTAACTCTTGATTCGTTAAACCCGCCAATCTCA 247
Qy      21      Phea1aGln1Ser1Gln1Pro1Gln1Leu1ySer1Gln1uSer1Val1le1Val1Ser1Arg1Gly 40
Db      248  TTGGTCTCAGAGTGAAGCCCGAGACTGAAGCTGGAAGAGTGTGTGATTTGTCAATGTCATGAT 307
Qy      41      Vala1rGa1aP1o1r1h1ySa1a1Thr1Gln1Leu1Met1Gln1Asp1Val1Thr1Pro1Sp1a1a1r1Pro 60
Db      308  GTGGGTGCTCAACCAAGGCCACCGAACTGATGACAGATGTCAACCCAGACGCAATGGCCA 367
Qy      61      Thr1r1r1Pro1Val1y1e1u1Gly1Gln1Leu1Thr1Pro1Arg1y1Gly1Gln1Leu1le1a1r1leu 80

```

```

Db      368 ACCGCGCCGGTAAACCTGGGTTGGCTGACACCGGNGGTGGTGAAGTAATCGCCTATCTC 427
QY      81  |||||TTPAAGLunAAGLValA1aAspGlyLeuLeuProLysCyseGlyCysePro 100
Db      428 GGAATTAATCTGGGCGACGGGTCTGGTGAACGAGGATGGCTGGCGAAAAAGGGCTCCCG 487
QY      101 GlnSerGlyGlnValA1aIleIleAAspValAAspGlnAAspGlnAAspGlnAAspGln 120
Db      488 CAGTCTGGTCAAGTCCCGATTAATGCTGATGTCAGAGCGTACCCGTAACAGGCGAA 547
QY      121 A1aApeA1aA1aGlyLeuA1aProAAspCyseA1aIleThrValA1aSerGlnA1aAspThr 140
Db      548 GCTTGGCGCGCGGCTGGCACTGACCTGGCAATACGTAATACCGAGAGATAG 607
QY      141 SerSerProAAspProLeuAAspAAspProLeuLysGlyValA1aCyseGlnLeuAAspA1a 160
Db      608 TCCAGTCCCGATCCGTTAATTAATCTTAACAACTGGGCTGGCCAACTGGATTAACGCG 667
QY      161 AAsnValThrAAspA1aIleLeuGlnAAspA1aGlySerIleA1aAAspPheThrGlyHis 180
Db      668 AACGTGACTGACCGCATCTCTACAGAGGCGAGAGGATCAATGCTAATCTTAACCGGCGCAT 727
QY      181 TyrgInThrA1aApeAAspGlnLeuGlnAAspValAAspAAspPheProGlnSerAAspLeuCy 200
Db      728 CGGCAAAACGGCGTTTGGCAACTGGAAACGGGTCTTAATTTTCCGCAATCAAACTTGTC 787
QY      201 LeuLysAAspGlnLysGlnAAspGlnLysGlnLysGlnLysGlnLysGlnLysGlnLys 220
Db      788 CTTAAACGTGAACAAACAGACAGAAAGCTGTTCAATTAACGACAGCATTAACATCGAATCTC 847
QY      221 LysValSerA1aAAspCyseValSerLeuThrGlnA1aValSerLeuA1aSerLeuThr 240
Db      848 AAGGTGAGGCGCGCAATCTCTTAACCGGTGGTGAAGCTTGGCAATCAAACTGAGAG 907
QY      241 GlnIlePheLeuLeuGlnGlnA1aGlnGlyMetProGlnProGlyTyrGlyAArgIleThr 260
Db      908 GAGATATTCTCTCTGCAACAGACAGAGGATCGCGAGACCGGGGTGGGAAAGATCAC 967
QY      261 AAspSerHisGlnIleThrAAspThrLeuLeuSerLeuHisAAsnA1aGlnPheAAspLeuGln 280
Db      968 GATTACACACAGATGAAACCTTGTCTAATGTTGCAATACCGCAATTTAATTTGCTACAA 1027
QY      281 ArgThrProGlnValA1aAAspSerAAspA1aThrProLeuLeuAAspLeuIleAAspThrA 300
Db      1028 CGCAGCGCCAGAGGTTGGCCGACGCGCACCCGCTTAATTAATTAATTAATTAATTAATTA 1087
QY      301 LeuThrProAAspProGlnLysGlnA1aTyrgIleValA1aThrLeuProThrSerValLeu 320
Db      1088 TTGACGCGCCCATCCACCGCAAAACAGGCGTATGGTGTGACATTAACCATTAAGTCTG 1147
QY      321 PheIleA1aGlyHisAAspThrAAspLeuA1aAAspLeuGlyValA1aLeuGlnLeuAAspThr 340
Db      1148 TTTATCGCGGACAGATTAATCTAATCTGGCAATCTCGCGGCGGCACTGGAGTCAACTGG 1207
QY      341 ThrLeuProGlnValA1aProAAspAAspThrProProGlnGlyLysLeuValA1aPheGlnAArg 360
Db      1208 ACGCTTCCCGGTCAACCGCATTAACGCGCGAGGTGGTGAACCTGGTGGTGAAGGCTGG 1267
QY      361 ArgArgLeuSerAAspAAspSerGlnIleGlnValSerLeuValA1aPheGlnThrLeuGln 380
Db      1268 CGTGGCTGAAGGATTAACAGGCAAGTGAATTCAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1327
QY      381 GlnMetAAspAAspLysThrProLeuSerLeuAAspThrProProGlnGlyValA1aLysLeuThr 400
Db      1328 CAGATGCGTGAATTAACCGCGCTGTCAATTAATTAATTAATTAATTAATTAATTAATTAAT 1387
QY      401 LeuA1aGlyCyseGlnLysAAspAAspA1aGlnGlyMetCyseSerLeuA1aGlyPheThrGln 420
Db      1388 CTGGAGAGATGTAAGAGAGCAAAATCGAGGCGAGGCAATGTTGTTGGCAGGTTTAACGCA 1447
QY      421 IleValAAspGlnA1aArgIleProA1aCyseSerLeu 432
Db      1448 ATCGTGAATGAAGCAGCATACCGGCTGCAAGTTTG 1483

```

```

RESULT 5
US-09-259-214-1
; Sequence 1, Application US/09259214A
; Patent No. 6110719
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHRYASE
; FILE REFERENCE: DIVER1370-1
; CURRENT APPLICATION NUMBER: US/09/259,214A
; CURRENT FILING DATE: 1999-03-01
; EARLIER APPLICATION NUMBER: 08/910,798
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; NAME/KEY: misc feature
; LOCATION: (1)...(1323)
; OTHER INFORMATION: n = A,T,C or G
US-09-259-214-1

Alignment Scores:
Pred. No.: 4,456-222 Length: 1323
Score: 2171.00 Matches: 422
Percent Similarity: 97.7% Conservative: 0
Best Local Similarity: 97.7% Mismatches: 10
Query Match: 96.1% Indels: 0
Gaps: 0

US-10-601-319-10 (1-432) x US-09-259-214-1 (1-1323)

QY      1  MetLysA1aIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerA1a 20
Db      1  ATGAAGGAGATCTTAATCCCATTTTATCTCTTCAATTCGGTAAACCCGCAATCTCA 60
QY      21  PheA1aGlnSerGlnProGlnLeuLysLeuGlnLysSerValA1aIleValSerHisGly 40
Db      61  TTGCTGAGAGTGAACCGGAGCTGAAGCTGAAGTGGTGAATGTCAGTCTCATGCT 120
QY      41  ValArgA1aProThrLysValA1aThrGlnLeuMetGlnAAspValA1aThrProAAspA1a 60
Db      121  GTGCTGCTCCAAACCAAGCCAGCACTGATGCAAGATGCAACCCAGAGCATGGCCA 180
QY      61  ThrTTPProValA1aLeuGlyGlnLeuThrProAArgIleGlyGlnLeuIleA1aTyrgLeu 80
Db      181  ACTGGCGCGTAAACCTGGTGGCTGACACCGGNGGTGGTGAAGTGAATGCGCTATCTC 240
QY      81  GlyHisTyrTTPAAGLunAAGLValA1aAAspGlyLeuLeuProLysCyseGlyCysePro 100
Db      241  GAACTTAACAAACCGCAGCGTCTGTGTAGCCAGATTTGCTGGCGAAAAAGGCTGCCG 300
QY      101  GlnSerGlyGlnValA1aIleIleAAspValAAspGlnAAspGlnAAspGlnAAspGln 120
Db      301  CAGTCTGGTCAAGTCCCGATTAATGCTGATGTCAGAGCGTACCCGTAACAGGCGAA 360
QY      121  A1aApeA1aA1aGlyLeuA1aProAAspCyseA1aIleThrValA1aSerGlnA1aAspThr 140
Db      361  GCTTGGCGCGCGGCTGGCACTGACCTGGCAATTAACCGTAATTAACCGAGATAGCG 420
QY      141  SerSerProAAspProLeuAAspAAspProLeuLysGlyValA1aCyseGlnLeuAAspA1a 160
Db      421  TCCAGTCCCGATCCGTTAATTAATCTTAACAACTGGGCTGGCCAACTGGATTAACGCG 480
QY      161  AAsnValThrAAspA1aIleLeuGlnAAspA1aGlySerIleA1aAAspPheThrGlyHis 180
Db      481  AACGTGACTGACCGGATCTCTACAGAGGCGAGAGGATCAATGCTGACTTAACCGGCGAT 540

```

```

Oy 181 TyrGlnThrAlaPheArgGluLeuGluuArgValLeuAenPheProGlnSerAsnLeuGys 200
    |||
Db 541 CGGCAACGGCGTTTCGGCAACTGGACGGGTGCTTAATTTCCGCATCAAACTTGTC 600
Oy 201 LeuYsArgGluYsGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
    |||
Db 601 CTTAAAGCTGAAGAACAGACGAAAGCTGTTCACTTAACGACGAGCTTACATCGGAATCC 660
Oy 221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
    |||
Db 661 AAGGTAGACGGCGCAATAGTCTCATTTAACCGGTGGTGAAGCTTCGACATGCTGACG 720
Oy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyValArgIleThr 260
    |||
Db 721 GAGATATTTCTCTGCAACAGACAGGGAATGCCGAGCCGGGTGGGGAAGATCAC 780
Oy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
    |||
Db 781 GATTACACACGATGGACACCTTGCTAAGTTTGACATAACGCGCAATTTTATTTGCTACA 840
Oy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleYsThrAla 300
    |||
Db 841 CGCAGCGCAGAGTTGCCCGCAGCGCGCACCCCGTTATTGATTTGATCATGGCAGCG 900
Oy 301 LeuThrProHisProProGlnuYsGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
    |||
Db 901 TTGACGCCCATCCACCGCAAAAACAGCGGCTGATGATGATGATTCACCTTCACTTCACTG 960
Oy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
    |||
Db 961 TTTATTTGCCGACACGATACCTAATCTGGCAATCTCGGGGGGCGCATGGAGCTCAACTGG 1020
Oy 341 ThrLeuProGlyValProAspAsnThrProProGlyGlyGluLeuValPheGluuArgTrp 360
    |||
Db 1021 ACGCTTCCCGGTACCGGATACACGCGCGCAAGGTGGAACTGCTGTTGAACGCTGG 1080
Oy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
    |||
Db 1081 CGTGGGCTTAAGGATTAACGACGATGATTCAGGTTTCCCTGCTCTTCCAGCTTACAG 1140
Oy 381 GlnMetArgAspYsThrProLeuSerLeuAsnThrProProGlyGlyValuYsLeuThr 400
    |||
Db 1141 CAGATGCGGTAAATAACCGCGCTGTCATTAATAACGCGCGCAGAGGTGAACCTGACC 1200
Oy 401 LeuAlaGlyCysGluGluuArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
    |||
Db 1201 CTGGCAGATGTGAAGAGCGGAATGCGCAGGCGCATGTGTCGTGGCAGGTTTACGCAA 1260
Oy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
    |||
Db 1261 ATCGTAATGAAGACGCAATACCGGCGTGCAGTTTG 1296

```

```

; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ... (1320)
; NAME/KEY: misc_feature
; LOCATION: (1) ... (1323)
; OTHER INFORMATION: n = A,T,C or G
US-09-318-528-1

Alignment Scores:
Pred. No.: 4,45e-222 Length: 1323
Score: 2171.00 Matches: 422
Percent Similarity: 97.7% Conservative: 0
Best Local Similarity: 97.7% Mismatches: 10
Query Match: 96.1% Indels: 0
DB: 3 Gaps: 0

US-10-601-319-10 (1-432) x US-09-318-528-1 (1-1323)
Oy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
    |||
Db 1 ATGAAAGCGATCTTAATCCATTTTATCTTCTGATTCGCTTAACCCGCAATCTGCA 60
Oy 21 PheAlaGlnSerGluProGluLeuYsGlnuSerValValIleValSerArgHisGly 40
    |||
Db 61 TTCCTCAGAGTGAAGCGGAGCTGAAGCTGAAGTGAAGTGGATGTCAGTCTCATGCT 120
Oy 41 ValArgAlaProThrLysValAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
    |||
Db 121 GTGGTGCTCCAAACCAAGGCCACGACGATGATGAGATGACCCAGACGACATGGCCA 180
Oy 61 ThrTrpProValYsLeuGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
    |||
Db 181 ACCTGGCCGGTAAATCGGTGGTGTGACACCGGNGTGGTGAATGCGCTATCTC 240
Oy 81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
    |||
Db 241 GGAATTAACCAACCCGACGCTGTGTAGCCGATTTGCTGGGAAAGGGGCTGCCCG 300
Oy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluuArgTrpArgYsThrGlyGlu 120
    |||
Db 301 CAGTGTGTGACGTGCGCATTAATGCTGATGATGATGATGATGATGATGATGATGATGAT 360
Oy 121 AlaPheAlaIleGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
    |||
Db 361 GCTTTCGCGCGCGGCTGGACCTGATCTGATTAACCGTACATCCAGGAGATACG 420
Oy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
    |||
Db 421 TCCAGTCCCGATCCGTTATTAACTCTTAATAAACTGGCGTTTGCAATGATTAACGCG 480
Oy 161 AsnValThrAspAlaIleLeuGluuArgAlaGlyYsSerIleAlaAspPheThrGlyHis 180
    |||
Db 481 AACGTGACTGACCGCATCTCAGACGAGGAGAGGATCAATGCTGACTTAACCGGGCAT 540
Oy 181 TyrGlnThrAlaPheArgGluLeuGluuArgValLeuAenPheProGlnSerAsnLeuGys 200
    |||
Db 541 CGGCAACGGCGTTTCGGCAACTGGACGGGTGCTTAATTTCCGCATCAAACTTGTC 600
Oy 201 LeuYsArgGluYsGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
    |||
Db 601 CTTAAAGCTGAAGAACAGACGAAAGCTGTTCACTTAACGACGAGCTTACATCGGAATCC 660
Oy 221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
    |||
Db 661 AAGGTAGACGGCGCAATAGTCTCATTTAACCGGTGGTGAAGCTTCGACATGCTGACG 720
Oy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyValArgIleThr 260
    |||
Db 721 GAGATATTTCTCTGCAACAGACAGGGAATGCCGAGCCGGGTGGGGAAGATCAC 780
Oy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
    |||

```

Db 781 GATTACACACAGTGAACACTTGTAGTTGCATTAACGCAATTTATTTGCTACAA 840
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleuValThrAla 300
Db 841 CGACCCCAAGAGTTCCCGACACCCGCGCACCCCGTTATTTGATTTGATGACGACG 900
Qy 301 LeuThrProHisProProGluIleuValArgIleValThrLeuProThrSerValLeu 320
Db 901 TTGAGCCCCCATCAACGCAAAAACAGGCGTATGCTGATGATTAACCACTTCACTAC 960
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeuGluLeuAsnTrp 340
Db 961 TTTATTTCCCGGACACCATATCTAATCTGGCAAAATCTGGCGGCGCACTGAGGCTCACTGG 1020
Qy 341 ThrLeuProGlyValInProAspAsnThrProProGlyValGluLeuValPheGluAsnTrp 360
Db 1021 AGCGTTCGGGTACGCGGATTAACAGCGCGGAGGTGAACTGGTGTGAAACCTGCG 1080
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1081 CGTGGCTTAAGCATTAACAGCCAGTGGATTGAGGTTTCGCTGCTTCCAGACTTTACAG 1140
Qy 381 GlnMetArgAspIleThrProLeuSerLeuAsnThrProProGlyValValIleuValLeuThr 400
Db 1141 CAGATCGTGAATAAACGCGCGTGTCTAATAATACCGCCCGGAGAGGTGAACCTGACC 1200
Qy 401 LeuAlaGlyCysGluGluValArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1201 CTGGCAGGATGTGAAGAGGAAATGCGCAGGGCATGTGTCGTGACAGTTTATACGAA 1260
Qy 421 IleValAsnGluValaArgIleProAlaCysSerLeu 432
Db 1261 ATCGTGAATGAACGACGCAATCCGCGCTGCAATTG 1296
RESULT 7
US-09-291-931-1
; Sequence 1, Application US/09291931A
; Patent No. 6190897
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/291,931A
; CURRENT FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; NAME/KEY: misc_feature
; LOCATION: (1)...(1323)
; OTHER INFORMATION: n = A,T,C or G
US-09-291-931-1
Alignment Scores:
Pred. No.: 4,45e-222 Length: 1323
Score: 2171.00 Matches: 422
Percent Similarity: 97.7% Conservative: 0
Best Local Similarity: 97.7% Mismatches: 10
Query Match: 96.1% Indels: 0
Db: 3 Gaps: 0
US-10-601-319-10 (1-432) x US-09-291-931-1 (1-1323)
Qy 1 MetValaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20

Db 1 ATGAAGGCAATCTTAATCCCAATTTTATCTTTCGATTCGGTAAACCCGCAATGTGCA 60
Qy 21 PheAlaGlnSerGluProGluLeuIleuValSerValIleValSerArgHisGly 40
Db 61 TTCGCTCAGAGTGAAGCCGAGAGCTGAAGCTGAAAGTGGTGAATGTCAAGTCATGGT 120
Qy 41 ValArgAlaProThrIleValaThrGluLeuMetGlnAspValaThrProAspAlaTrpPro 60
Db 121 GTGCGTGTCCCAACCAAGGCAACGCAATGAGAGAGTGAACCCGCAACGATGCGCA 180
Qy 61 ThrTrpProValIleuLeuGlyValLeuThrProArgGlyGlyValLeuIleAlaTyrlLeu 80
Db 181 ACCTGGCCGGTAAACTGGGTGGCTGACACCCGAGGTGGTGAAGTAAATCCGCTATCTC 240
Qy 81 GlyHisTyrlTrpArgGlnArgLeuValAlaAspGlyLeuLeuProIleCysGlyCysPro 100
Db 241 GGACATTAACCAACGCAAGCGTCTGGTAAAGCGAGTTGCTGGCGMAAAGGCGTCCCG 300
Qy 101 GlnSerGlyValIleValIleIleIleAlaAspValaAspGluArgThrArgIleThrGlyIle 120
Db 301 CAGTCTGTCAAGTGGCGATTAATGCTGATGTGCAAGACCGTAAACCGCGTAAACAGCGCA 360
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCTTCGCGCGCGGCTGGCACTTAAGTGTGAATTAACCGTACATTAACGAGGCAATACG 420
Qy 141 SerSerProAspProLeuPheAsnProLeuIleuValSerArgIleValaCysGluLeuAspAsnAla 160
Db 421 TCCAGTCCCGGATCCGTTATTTAATCCTGTAATAATCGGGGTTTGGCAATGATGAACGCG 480
Qy 161 AsnValaThrAspAlaIleLeuGluValArgAlaGlyIleSerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGACTGAAGCCATCTCAAGCAGGCGAGGAGGTCAATTCGACTTTTACCGGCGCAT 540
Qy 181 TyrlTrpThrAlaPheArgGluLeuGluValArgValaLeuAsnProGlnSerAsnLeuCys 200
Db 541 CGGCAAAACGCGCTTTCGGAACTGAACGGGAGCTTAATTTTCGCAATCAAACTGTGGC 600
Qy 201 LeuIleValArgIleValaAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAAACGTGAAGAAACAGAGCAAGAACTGTTCATTAACGAGGCAATTACATCGGAATCC 660
Qy 221 LysValSerAlaAspCysValSerLeuThrGlyValaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGAACGCGCGAATGTCTCATTAACCGGTGGGTGAAGCTGCATCAATGCTGACG 720
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrlGlyValArgIleThr 260
Db 721 GAGATATTTCTCTCAACAGCAACAGGAAATGCGGAGCGGGGTGGGAAAGATCAAC 780
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
Db 781 GATTCAACCAAGTGAACACCTGTGTAAGTTTGGCAATTAACGCGCATTTTATTTGCTACA 840
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleuValThrAla 300
Db 841 CGACGCGCAGAGTGGCCCGGACGCGCGCAACCCCGTTATTTGATATGATCATGCGACG 900
Qy 301 LeuThrProHisProProGluIleuValArgIleValaThrLeuProThrSerValLeu 320
Db 901 TTGAGCCCCCATCAACGCAAAAACAGGCGTATGCTGATGATTAACCACTTCACTAC 960
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeuGluLeuAsnTrp 340
Db 961 TTTATTTCCCGGACACCATATCTAATCTGGCAAAATCTGGCGGCGCACTGAGGCTCACTGG 1020
Qy 341 ThrLeuProGlyValInProAspAsnThrProProGlyValGluLeuValPheGluAsnTrp 360
Db 1021 AGCGTTCGGGTACGCGGATTAACAGCGCGGAGGTGTAAGTGTGTAACGCTGGAACGCTGG 1080
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380

Db 1081 CGTCGGCTAAGCATTAACGACGAGTACGATTTCGTGCTTCAGACTTTACAG 1140
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGluValLeuThr 400
Db 1141 CAGATGCGTGAATAAAGCCCGCTGCTAATAATACGCCCGGAGAGTAAACTGACC 1200
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1201 CTGGAGAGATGTGAAGACGAAATCCGACGGGCAATGTTCTTGCGACGATTTAACGAA 1260
Qy 421 lIeValaengluAlaArgIleProAlaCysSerLeu 432
Db 1261 ATCGTAATGAAGCAGCATACCGCGCTGCAGATTG 1296
RESULT 8
US-09-580-515-1
; Sequence 1, Application US/09580515
; Patent No. 6720014
; GENERAL INFORMATION:
; APPLICANT: Kreitz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/580, 515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; NAME/KEY: mIsC_feature
; LOCATION: (1)...(1323)
; OTHER INFORMATION: n = A,T,C or G
US-09-580-515-1
Alignment Scores:
Pred. No.: 4,45e-222 Length: 1323
Score: 2171.00 Matches: 422
Percent Similarity: 97.7% Conservative: 0
Best Local Similarity: 97.7% Mismatches: 10
Query Match: 96.1% Indels: 0
DB: 3 Gaps: 0
US-10-601-319-10 (1-432) x US-09-580-515-1 (1-1323)
Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAAGGAGATTTAATCCCATTTTATCTCTTCGATTCGGTTAACCCCGCAATCTGCA 60
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleIleValSerArgHisGly 40
Db 61 TTCGCTCAGAGTGAAGCCGAGCTGAAGCTGGAAGAGTGTGATGTGATGTGATGTGATGT 120
Qy 41 ValAlaGluAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 121 GTCGCTGTCTCCAAACGAGCCACGCAACTGATGACAGATGTCACCCGACGATGCGCA 180
Qy 61 ThrTrpProValLysLeuGluGluLeuThrProArgGluGluGluLeuIleLeuIleLeu 80
Db 181 ACCTGGCCGGTAAACCTGGGTTGGCTGACACCGCAGGGTGTGAGCTAATCCCTATCTC 240
Qy 81 GlyHisTyrtTrpArgGlnArgLeuValAlaSerGlyLeuLeuProLysCysGlyCysPro 100

Db 241 GGACATTACCAAGCCGACGCTGTGTAGCCGACGGAATTGCTGGCGGAAAAAGGCTGCCCG 300
Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgTrpArgLysThrGlyGlu 120
Db 301 CAGTCTGTGAGGTGGGATTAATGCTGATGTCAACAGGGTACCCGTAAACAGGCGAA 360
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCTTCGCCCGCGGGCTGGCACCCTGACTGTGCATTAACCGTACCTACCGGACGATACG 420
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAla 160
Db 421 TCCAGTCCCGATCCGTTATTTAATCTCTAAACCTGGCGTTGGCCAACTGATTAACGCG 480
Qy 161 AsnValThrAspAlaIleLeuGluArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGACTGACGGGATCTTCAGAGGGCAGAGAGTCAATGTGCTACTTTAACCGGCGAT 540
Qy 181 TyrGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 541 CGGAAACGGCGTTTCCCGAATCGGAACGGGTCTTAATTTCCGCAATCAAACTTGTC 600
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAACGTTGAGAAACAGACAGAAAGCTGTTCATTAAACGACGACATTACCATCGAACTC 660
Qy 221 LysValSerAlaAspCysValSerLeuThrGlyValAlaSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGAGCGCGCAATGTCTCAATTAACCGGTCCGGTAAGCTCGCATCAATCTGACG 720
Qy 241 GluIlePheLeuLeuGlnAlaGlnAlaGlnGlyMetProGluProGlyTrpArgLysThr 260
Db 721 GAGATATTTCTCCGCAACAGACAGGAAATGCCGAGACCGGGGTGGGAAGATCACCC 780
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
Db 781 GATTCAACACAGTGAACACCTGCTTAAGTTTGATTAACCGGCAATTTTATTCCTCAAA 840
Qy 281 ArgThrProGluValAlaIleArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
Db 841 CGCAGCCGACAGAGTTGCCCGGACCGCGCACCCCGTTATTTGACATTGATCATGACGCG 900
Qy 301 LeuThrProHisProProGlnLysGlnAlaIleArgLysValThrLeuProThrSerValLeu 320
Db 901 TTGAAGCCCATCAACCGCAAAAACAGCGATATGTGTGACATTACCACTTGATGTCGTG 960
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 961 TTTATTTCCCGACACGATTAATCTGGCAATCTCGCGCGCGACCTGAGCTCAACTCG 1020
Qy 341 ThrLeuProGluGluProAspAsnThrProProGluGlyGluLeuValPheGluArgTrp 360
Db 1021 ACGCTTCCCGGTACCGGATTAACACCGCGCAGGTGTGTGAACCTGTGTGAACGCTGG 1080
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1081 CGTCGGCTAAGCATTAACGACGACGATTAAGTTTCGCTTCGCTTCAGACTTTACAG 1140
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGluValLeuThr 400
Db 1141 CAGATGCGTGAATAAAGCCCGCTGCTAATAATCGCCGCGGAGAGGTAAACTGACC 1200
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1201 CTGGAGAGATGTGAAGACGAAATGCGAGGGCAATGTTTGTGTGACAGGTTTAACGAA 1260
Qy 421 lIeValaengluAlaArgIleProAlaCysSerLeu 432
Db 1261 ATCGTAATGAAGCAGCATACCGCGCTGCAGATTG 1296
RESULT 9
US-09-586-379-1

```
/ Sequence 1, Application US/09866379
/ Patent No. 6855365
/ GENERAL INFORMATION:
/ APPLICANT: DIVERSA CORPORATION
/ APPLICANT: SHORT, Jay
/ APPLICANT: KREITZ, Keith
/ APPLICANT: GRAY, Kevin
/ APPLICANT: BARTON, Nelson
/ APPLICANT: GARRETT, James
/ APPLICANT: O'DONOGHUE, Eileen
/ TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
/ FILE REFERENCE: DIVER1370-7
/ CURRENT APPLICATION NUMBER: US/09/866,379
/ PRIOR FILING DATE: 2001-05-24
/ PRIOR APPLICATION NUMBER: US 09/560,515
/ PRIOR FILING DATE: 2000-05-25
/ PRIOR APPLICATION NUMBER: US 09/318,528
/ PRIOR FILING DATE: 1999-05-25
/ PRIOR APPLICATION NUMBER: US 09/291,931
/ PRIOR FILING DATE: 1999-04-13
/ PRIOR APPLICATION NUMBER: US 09/259,214
/ PRIOR FILING DATE: 1999-03-01
/ PRIOR APPLICATION NUMBER: US 08/910,798
/ PRIOR FILING DATE: 1997-08-13
/ NUMBER OF SEQ ID NOS: 10
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 1
/ LENGTH: 1323
/ TYPE: DNA
/ ORGANISM: Escherichia coli
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)..(1323)
/ OTHER INFORMATION: n is any nucleotide
/ NAME/KEY: CDS
/ LOCATION: (1)..(1323)
/ OTHER INFORMATION:
/ US-09-866-379-1

Alignment Scores:
Pred. No.: 4,45e-222 Length: 1323
Score: 2171.00 Matches: 422
Percent Similarity: 97.7% Conservative: 0
Best Local Similarity: 97.7% Mismatches: 10
Query Match: 96.1% Indels: 0
Gaps: 0
DB: 3

US-10-601-319-10 (1-432) x US-09-866-379-1 (1-1323)

QY 1 MetLysAlaIleuLleProPhleuSerLeuLleProLeuThProGInSerAla 20
DB 1 ATGAAAGCATCTTAATCCCATTTTATCTCTGATTCCGTTAACCCCGCAATCTGCA 60
QY 21 PheAlaGInSerGluProGluLeuLysLeuGInSerValIleValSerArgHisGly 40
DB 61 TTCGCTCAGAGGAGCCGAGCTGAAGCTGAAAGTGAGTATGTCAGTGTCAATGCT 120
QY 41 ValAlaGInProThrLysAlaThrGInLeuMetGInAspValThrProAspAlaTrpPro 60
DB 121 GTGGCTGCTCCAAACGAGCCAGCACTGATGCAAGATGTCACCCCAACCATGGCCA 180
QY 61 ThrTrpProValLysLeuGluLysLeuLysProArgGlyGlyLysLeuLleAlaTrpLeu 80
DB 181 ACCTGCGCGTAAACCTGGGTGGTGGCTGACACCGCAGGAGTGAGCTAAATCCCTATCTC 240
QY 81 GlyHisTrpTrpArgGInArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
DB 241 GGAACATTACCAACGCGCAAGGCTGTGGTGAAGCGACGATTCCTGCGAAAAAGGCTGCCG 300
QY 101 GInSerGlyGInValAlaIleIleAlaAspValAspGluArgTrpArgLysThrGlyGlu 120
DB 301 CAGTCTGTGACAGTGGCGATTAATGCTGATGTGACGAGCGTAACCGTTAAACAGGCGCA 360
```

```
QY 121 AlaPheAlaIleGluLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
DB 361 GCCCTGCGCGCGGCTGGCACCCTGACGTGCAATAACGTTACATACCCAGGAGATAG 420
QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGInLeuAspAla 160
DB 421 TCACGTCCGATCCGTTATTTATCTTAATCCTTAATAACGCGGCTTCCAACTGATACCGG 480
QY 161 AsnValThrAspAlaIleLeuGluLysArgAlaGlySerIleAlaAspThrGlyHis 180
DB 481 AACGTGACTGACGCGCATCTTCAGAGGCGAGGAGGTCAATTGCTGACTTTACCGGCGAT 540
QY 181 TrpGInThrAlaPheArgGluLeuGluLysValLysAsnPheProGInSerLeuLysCys 200
DB 541 CCGCAACCGCGCTTCGCAACCTGAAACGGGTGCTTAATTTCCGCAATCAACCTGTGC 600
QY 201 LeuLysArgGluLysGInAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
DB 601 CTTAAACGTGAAGAACAGACCAAAAGCTGTTCAATTACGACGACATTACATCGAACTC 660
QY 221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
DB 661 AAGTGAGCGCGCAAAATGCTCATTAACCGGTGCGTTAGCTCGCATCAATGCTGACG 720
QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
DB 721 GAGATATTTCTCTGCAACACAGGAAATGCGGAGCGGGGTGGGAAAGATCACCC 780
QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
DB 781 GATTCACACACAGTGAACACCTTGCTAGTTGCTAATACGCCCAATTTATTTGCTACAA 840
QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
DB 841 CGCAGCGCAAGAGGTGGCGCGCAGCGCGCACCCGTTATGATGATTCATGACAGCGG 900
QY 301 LeuThrProHisLeuProProGluLysGlnAlaTrpGlyValThrLeuProThrserValLeu 320
DB 901 TTGACGCCCATCCACCCCAAAACAGCGATGATGTGATGATCCCACTTCAGTACATCG 960
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyValAlaLeuGluLeuAsnTrp 340
DB 961 TTTATTTGCGGACACAGATACATTAATTTGGCAATCTCGCGCGCGCATCTGAACTCAACTGG 1020
QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
DB 1021 ACGCTTCCCGGTCAACCGGATTAACACGCGCGAGGTGGTGAACGTGTGTTGAACGCTGG 1080
QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
DB 1081 CGTCCGCTTAAGGATTAACAGCCAGTGAATTCAGGTTTGGCTGCTCCAGACTTTACAG 1140
QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
DB 1141 CAGATCGCTGATTAACCCCGCTGCTCATTAATACCGCCCGGAGAGGTGAATCTGACC 1200
QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
DB 1201 CTGGCAGGATGTGAAGACGAAATGCGCAGGGCATGTGTTGTTGGTCAGAGTTTACGCA 1260
QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
DB 1261 ATCGTAATGAAGCAGCATACCGCGCTGCAAGTTTG 1296

RESULT 10
US-09-715-477-2
/ Sequence 2, Application US/09715477
/ Patent No. 6841370
/ GENERAL INFORMATION:
/ APPLICANT: Lei, Xingan
/ TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYTASE
/ FILE REFERENCE: 19603/4031
/ CURRENT APPLICATION NUMBER: US/09/715,477
```

```

; CURRENT FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,179
; PRIOR FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1486
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-715-477-2

Alignment Scores:
Pred. No.: 3e-221 Length: 1486
Score: 2164.00 Matches: 421
Percent Similarity: 97.5% Conservative: 0
Best Local Similarity: 97.5% Mismatches: 11
Query Match: 95.8% Indels: 0
DB: Gaps: 0

US-10-601-319-10 (1-432) x US-09-715-477-2 (1-1486)

QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
DB 188 ATGAAGGAGATCTTATCCCATTTTATCTTCTGATTCGGTTAACCCCGCAATCTGCA 247
QY 21 PheAlaGlnSerGluProGlnLeuLysLeuGlnSerValIleValSerArgHisGly 40
DB 248 TTCGCTCAGATGACCCGAGCTGAAGCTGAAAGTGTGGTATGTGACCGCTCATGCT 307
QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
DB 308 GTGGCTGCCCAACCAAGGCCGACGCACTGATGACGATGTCACCCGACGCAATGGCCA 367
QY 61 ThrTrpProValLysLeuGlyLysLeuThrProArgGlyGlyGlnLeuIleAlaTyrLeu 80
DB 368 AACTGGCCGGTAAACCTGGGTGGCTGACACCAACCGGTGGAGCTAATGCCATATCTC 427
QY 81 GlyHisTyrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
DB 428 GGACATTACCAACGCGACGCTGTGGTGGCCGACGATTTGCTGGCAAAAAGGCTGGCCG 487
QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgLysThrGlyGln 120
DB 488 CAGCTGTGTCAGTGTGGCATTTTGTGATGTCGACGAGCGTAAACCGTAAACAGGCCAA 547
QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
DB 548 GCGTTCGCGCGCGGCTGGCACTGATCTGCAATTAACCATTAACCCAGGCAATACG 607
QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
DB 608 TTCAGTCCCGATCCCTTATTTTCTCTTAATAACCTGGGCTTGGCAACTGATTAACGG 667
QY 161 AsnValThrAspAlaIleLeuGlnLysArgLysSerIleAlaAspPheThrGlyHis 180
DB 668 AACGTGACTGACGCGATCCTCAAGGCGAGGAGGTCAATTGCTTATCCGGGCAT 727
QY 181 TyrGlnThrAlaPheArgGlnLeuGlnLysArgValIleAsnPheProGlnSerAsnLys 200
DB 728 CGGCAAAAGCGGCTTTCGCAACTGAAACGGGTGCTTAATTTTCCCAATCAAACTGTGTC 787
QY 201 LeuLysArgGlnLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlnLeu 220
DB 788 CTTAAACGTGAAACAGGACGAAAGCTGTTTCAATTACGACGACATTACATCGAATCTC 847
QY 221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
DB 848 AAGGTAGAGGCGCGCAATGTTTCAATTAAACGGTGGGTAAAGCTCGCATCAATGTGACG 907
QY 241 GlnIlePheLeuLeuGlnAlaGlnGlyMetProGlnProGlyTyrGlyArgGlyLeuThr 260
DB 908 GAAATATTTTCTCTCCACACAGCACAGGAAATGCGGAGCCGGGTGGGGAAGGATCACT 967
```

```

QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuGln 280
DB 968 GATTACACACGAGTGAACACCTTCTAAGTTTGATATACCGCAATTTTATTACTACAA 1027
QY 281 ArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
DB 1028 CGCAGCGCAGAGGTGTGCGCGGACGTGCGCCACCCCTTAATGTGATTCAGACAGACGCG 1087
QY 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
DB 1088 TTGAGGCCCATCCACCGCAAAACAGCGGATGATGTGATTAACCATTCCTGAGTGTG 1147
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTrp 340
DB 1148 TTTATTTGCCGACACGATACATCAATCTGCGGCGCGCACCTGAGACTCAACTCGG 1207
QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlnLeuValPheGlnArgTrp 360
DB 1208 ACGCTTCAGGTGACCGCGATTAACACGCGCGCAGGTGTGTAACTGTGTGTGAACGCTGG 1267
QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
DB 1268 GGTGGCTAAAGCGATTAACAGCAAGTGAATCAGGTTCGCTGTCTTCACACTTAAAG 1327
QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlnValLysLeuThr 400
DB 1328 CAGATGGGTGATTAACGCGCTATCATTTAAATACGCGCGCGAGAGGTGAACCTGACC 1387
QY 401 LeuAlaGlyCysGlnGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
DB 1388 CTGGCAGAGATGTGAAGAGCGAAATGCGCAGGCGCATGTGTGTCGTGGCGGTTTACCAA 1447
QY 421 IleValAsnGlnAlaArgGlyLeuProAlaCysSerLeu 432
DB 1448 ATCGTAATGAAGCGCGATACCGGCTGCAgTTTG 1483

RESULT 11
US-09-540-149A-9
; Sequence 9, Application US/09540149A
; Patent No. 6511699
; GENERAL INFORMATION:
; APPLICANT: Lei, Xiangen
; TITLE OF INVENTION: ENZYMES WITH IMPROVED PHYTASE ACTIVITY
; FILE REFERENCE: 19603/2791
; CURRENT APPLICATION NUMBER: US/09/540,149A
; CURRENT FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 60/127,032
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1489
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-540-149A-9

Alignment Scores:
Pred. No.: 1.98e-219 Length: 1489
Score: 2147.00 Matches: 418
Percent Similarity: 96.8% Conservative: 0
Best Local Similarity: 96.8% Mismatches: 14
Query Match: 95.1% Indels: 0
DB: Gaps: 0

US-10-601-319-10 (1-432) x US-09-540-149A-9 (1-1489)

QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
DB 182 ATGAAGGAGATCTTATCCCATTTTATCTTCTGATTCGGTTAACCCCGCAATCTGCA 241
QY 21 PheAlaGlnSerGluProGlnLeuLysLeuGlnSerValIleValSerArgHisGly 40
DB 242 TTCGCTCAGATGACCCGAGCTGAAGCTGAAAGTGTGGTATGTGACCGCTCATGCT 301
```

```

QY      41 ValArgAlaProThrIysAlaThrGlnLeuMetGlnAspValThrProAspAlaTTPPro 60
DB      302 GTGGCGGCCCAACAAAGCCACCAATGATGACAGATGTCACCCCAAGCGATGCGCA 361
QY      61 ThrTPProValIlyLeuGlyGlnLeuThrProArgGlyGlyGlnLeuIleAlaTyrLeu 80
DB      362 ACCGTGCGCGGTAAACCTGGGTTGGCTGACACACGCGGTGGTGAAGCTAATCGCTATCTC 421
QY      81 GlyHisTyrTTPArgGlnIleuValAlaAspGlyLeuLeuProIysCysGlyCysPro 100
DB      422 GGACATATTCACACGCGCGTCTGGCGCACGAGATTCGTGGCAAAAAGGCGCTCGCCG 481
QY      101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGlnArgThrArgIysThrGlyGln 120
DB      482 CAGCTGCTCAGGTCCGCAATATGCTGATGACAGCGTAAACCGGTAAACAGGCGCA 541
QY      121 AlaPheAlaAlaGlyLeuValaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
DB      542 GCCTTCGCGCGCGGCTGGCACTGACCTGTGCAATACCTGACATACCCAGGCAAGTACG 601
QY      141 SerSerProAspProLeuPheAsnProLeuIysThrGlyValCysGlnLeuAspAsnAla 160
DB      602 TCCAGTCCCGATCCGTTATTTAATCCTTAAACCTGGCGTTGCGCACTGGATACGCG 661
QY      161 AsnValThrAspAlaIleLeuGlnArgAlaGlyIysSerIleAlaAspPheThrGlyHis 180
DB      662 AACGTGACAGCGCATCTCTCAGCAGGCGAGGCGTCAATGCTCACTTACCGCGGCAT 721
QY      181 TyrGlnThrAlaPheArgGlyLeuGlnArgValIleuAsnPheProGlnSerAsnLeuCys 200
DB      722 CGGCAACGCGCTTTCGCAACTGCAACGCGGTGCTTAATTTTCCCAATTAACCTTGTC 781
QY      201 LeuIysArgGlnIlyGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlnLeu 220
DB      782 CTTAACCGTGAGAAACAGGACGAAAGCTGTTCATTACACGAGCATTAACATCGGAATCTC 841
QY      221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
DB      842 AAGGTAGAGCGCGCAATGTTTCATTACCGGTGCGGTAAAGCTGCAATTAATGCTGACG 901
QY      241 GlnIlePheLeuLeuGlnIleAlaGlnIlyMetProGlnPProGlyTTPGlyValIleThr 260
DB      902 GAAATATTTCTCTCTGCAACACACAGGAAATGCGGAGCCGCGGTGGGAAAGCATCTCT 961
QY      261 AspSerHisGlnTTPAsnThrIleuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
DB      962 GATTCAACACAGTGAACACCTTGTCAAGTTTGCATTAACGGCAATTTTATTACTACAA 1021
QY      281 ArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIleIysThrAla 300
DB      1022 CGCAGCGCAAGAGTTGCCGCGAGTGGCGCACCCCGTTATGGAATTTGATCAATGCGACG 1081
QY      301 LeuThrProHisAspProGlnIlyGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
DB      1082 TTGAGCGCCCATCCACCGCAAAAACAGCGGTGATGATGACATTAACCACTTCAATGCTG 1141
QY      321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnThr 340
DB      1142 TTTATTTCCGACACGATACCTAATCTGGCAATCTCGGCGGCGGCACTGAGCTCAACTG 1201
QY      341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlnLeuValPheGlnArgTrp 360
DB      1202 ACGCTTCAAGGTCAACCGGATTAACAGCGCGCAGGTGTGTGAACCTGTGTTGAAGCGTGG 1261
QY      361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
DB      1262 CGTCGGCTTAAGGATTAACGCAAGTGAATTCAGGTTTGGCTGGTCTTCCAGACTTACAG 1321
QY      381 GlnMetArgAspIlyThrProLeuSerLeuAsnThrProProGlyGlyValIlyLeuThr 400
DB      1322 CAGATGCGTGAATAAAACCGCGTATCAATTAATAACGCGCGCGGAGAGGTGAATACGAC 1381

```

```

QY      401 LeuAlaGlyCysGlnGlnArgAsnAlaGlnIlyMetCysSerLeuAlaGlyPheThrGln 420
DB      1382 CTGGCAGATGTGAAGAGCGAAATGCGCAGGCGCATGTTGCTGCGGTTTACGCA 1441
QY      421 IleValAsnGlnAlaArgIleProAlaCysSerLeu 432
DB      1442 ATCGTGAATGAAGCGCGCATACCGCGCTGCATTTG 1477

RESULT 12
US-10-266-041A-9
; Sequence 9, Application US/10266041A
; Patent No. 6974690
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingren
; TITLE OF INVENTION: PHOSPHATASES WITH IMPROVED PHYTASE ACTIVITY
; FILE REFERENCE: 19603/2798
; CURRENT APPLICATION NUMBER: US/10/266,041A
; PRIOR FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: 60/127,032
; PRIOR FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 09/540,149
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1489
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-266-041A-9

Alignment Scores:
Pred. No.: 1,986-219 Length: 1489
Score: 2147.00 Matches: 418
Percent Similarity: 96.84 Conservative: 0
Best Local Similarity: 96.84 Mismatches: 14
Query Match: 95.14 Indels: 0
DB: 4 Gaps: 0

US-10-601-319-10 (1-432) x US-10-266-041A-9 (1-1489)
QY      1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
DB      182 ATGAAAGCATCTTAATCCATTTTATCTTTGATTCGGTAAACCCGCAATCTGCA 241
QY      21 PheAlaGlnSerGlnProGlnLeuIysLeuGlnSerValIleValSerArgHisGly 40
DB      242 TTCGCTCAGATGAGCGCGGAGCTGAAGCTGGAAGTGTGGTGAATGTCAACCGCTATGCT 301
QY      41 ValArgAlaProThrIysValaThrGlnLeuMetGlnAspValThrProAspAlaTTPPro 60
DB      302 GTGGCGGCCCAACAAAGCCACCAATGATGACAGATGTCACCCCAAGCGATGCGCA 361
QY      61 ThrTPProValIlyLeuGlyGlnLeuThrProArgGlyGlyGlnLeuIleAlaTyrLeu 80
DB      362 ACCGTGCGCGGTAAACCTGGGTTGGCTGACACACGCGGTGGTGAAGCTAATCGCTATCTC 421
QY      81 GlyHisTyrTTPArgGlnIleuValAlaAspGlyLeuLeuProIysCysGlyCysPro 100
DB      422 GGACATATTCACACGCGCGTCTGGCGCACGAGATTCGTGGCAAAAAGGCGCTCGCCG 481
QY      101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGlnArgThrArgIysThrGlyGln 120
DB      482 CAGCTGCTCAGGTCCGCAATATGCTGATGACAGCGTAAACCGGTAAACAGGCGCA 541
QY      121 AlaPheAlaAlaGlyLeuValaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
DB      542 GCCTTCGCGCGCGGCTGGCACTGACCTGTGCAATACCTGACATACCCAGGCAAGTACG 601
QY      141 SerSerProAspProLeuPheAsnProLeuIysThrGlyValCysGlnLeuAspAsnAla 160
DB      602 TCCAGTCCCGATCCGTTATTTAATCCTTAAACCTGGCGTTTCCAACTGGATACGCG 661
QY      161 AsnValThrAspAlaIleLeuGlnArgAlaGlyIysSerIleAlaAspPheThrGlyHis 180

```

```
Db 662 AACGGACTGACGGGATCTCAGACAGGAGGAGTCAATTGCGACTTTACCGGCGAT 721
Qy 181 TrrglnThraAlaPheArgGluLeuGluValLeuAsnProGlnSerAsnLeuCys 200
Db 722 CGGAAACGGGGTTCGGAACTGGAAACGGGGTCTTAATTTTCCCAATTAACTTGTC 781
Qy 201 LeuYsaArgGluYsglnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 782 CTTAACCGTGAAGAAACAGACGAAAGCTGTTCACTTAACGAGGCACTTCCATCGAATC 841
Qy 221 LysValSerAlaAspCysValSerLeuThrGlnAlaValSerLeuAlaSerMetLeuThr 240
Db 842 AAGGTAGACGGCCGAATGTTTCACTTAACCGGTGGTAAGCTCGATCATGTCAGCG 901
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyValArgIleThr 260
Db 902 GAAATATTCTTCCTGGCAACAAGCAACGGAAATGCCGAGCCGGGTGGGGAAGATCACT 961
Qy 261 AspSerHisGlnTrrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
Db 962 GATTCAACCAAGTGAACACCTTGCTTAAGTTTGATTAACGCGCAATTTTATTACTCA 1021
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleYsrThrAla 300
Db 1022 CGCAGCGCAGAGGTGGCCGCGAGTCGCCACCCCGTTATGGATTTGATCATGGCAGCG 1081
Qy 301 LeuThrProHisProProGlnIlysglnAlaTyrGlyValIleThrLeuProThrSerValLeu 320
Db 1082 TTGAGCGCCCATCCACCCGCAAAAACAGGCGTATGTGTGATTAACCACTTCAAGTGTG 1141
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluIleAsnThr 340
Db 1142 TTTATTCGCGACACGATACTATATGCGAAATCTCGCGCGCGCACTGTAGCTCACTCG 1201
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrrp 360
Db 1202 ACGCTTCAGGTCAGCCGGATTAACACGCCGCCAGGTGTGAATCTGTGTGAACGCTGG 1261
Qy 361 ArgArgLeuSerAspAsnSerGlnTrrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1262 CGTGGCTTAACCGATTAACAGCCAGTGAATTCAGGTTTGGCTGCTTCACACTTTACAG 1321
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValIleYsrThr 400
Db 1322 CAGATGCGTGAATAACCCGCTATCATTAATACGCCGCCGAGAGGTAACTGACC 1381
Qy 401 LeuAlaGlyCysGluGluValArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1382 CTGGCAGGATGTAAGACGAAATCGCAGGGGCATGTGTGTTGGCCGGTTTACGCA 1441
Qy 421 IleValaGlnGluAlaArgIleProAlaCysSerLeu 432
Db 1442 ATCGTAATGAAGCGCATACCGCGTGGCAGTTTG 1477

RESULT 13
US-09-715-477-4
; Sequence 4, Application US/09715477
; Patent No. 6841370
; GENERAL INFORMATION:
; APPLICANT: lei, Xingren
; TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYTASE
; FILE REFERENCE: 19603/4031
; CURRENT APPLICATION NUMBER: US/09/715,477
; CURRENT FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,179
; PRIOR FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1486
; TYPE: DNA
; ORGANISM: Escherichia coli
```

```
US-09-715-477-4
Alignment Scores:
Pred. No.: 4,13e-219 Length: 1486
Score: 2144.00 Matches: 418
Percent Similarity: 97.2% Conservative: 2
Best Local Similarity: 96.8% Mismatches: 12
Query Match: 95.0% Indels: 0
DB: 3 Gaps: 0

US-10-601-319-10 (1-432) x US-09-715-477-4 (1-1486)
Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 188 ATGAAACGATCTTAATCCATTTTATCTTGTGATTCGTTAACCAGGCAATCTGCA 247
Qy 21 PheAlaGlnSerGluProGluLeuYsglnLeuGlnSerValIleValSerArgHisGly 40
Db 248 TTGCGTCAAGTGGAGCCGAGCTGAAGCTGAAAGTGTGGATGTGCACCGCTCATGCT 307
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValIleThrProAspAlaTrrpPro 60
Db 308 GTGCTGCCCCAACCAAGGCCAGCAACTGATGCAAGATGTCAACCCAGACGAGTGGCA 367
Qy 61 ThrTrrpProValIlysglnLeuGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db 368 ACCGTGCGGGTAAAACTGGGTGGCTGACACACGCGGTGGTGAATTCGCTATCTC 427
Qy 81 GlyHisTrrpTrrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
Db 428 GGAATTATCAACAGCCCGCTGTGGTGGCCGAGGATTTGCTGGCGAAAAAGGCGTCCCG 487
Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgTrrpArgLysThrGlyGlu 120
Db 488 CAGCTGTCAAGTTCGCGATTTATGTCATGTGCACAGCGCTACCCGTAACACAGCGAA 547
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCCCTGCGCGCGGGCTGGACCTGACTGTGCAATTAACCGTAACACGAGATACG 607
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 608 TCCAGTCCCGATCCGTTATTTATTCCTCTAAACCTGCGCTTGGCCAACTGATMAACGCG 667
Qy 161 AsnValIleAspAlaIleLeuGluArgAlaGlyIleSerIleAlaAspPheThrGlyHis 180
Db 668 AACGTGACTGACGGGATCTCAGCAGGCGAGAGGGTCAATGCTGACTTTACCGGCGAT 727
Qy 181 TrrglnThraAlaPheArgGluLeuGluValLeuAsnProGlnSerAsnLeuCys 200
Db 728 CGGAAACGGGGTTCGGCAACTGGAACGGGTGCTTAATTTTCCGAATCAAACTTGAC 787
Qy 201 LeuYsaArgGluYsglnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 788 CTTAAACGTGAAGAAACGAATGAAGCTGTAACTTAAACGAGCATTAACATCGAATC 847
Qy 221 LysValSerAlaAspCysValSerLeuThrGlnAlaValSerLeuAlaSerMetLeuThr 240
Db 848 AAGGTAGACGGCCGAATGTTTCACTTAACCGGTGGTAAGCTCGATCAATCTAGCG 907
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyValArgIleThr 260
Db 908 GAAATATTCTTCGGAACAAGCAACAGGAATGCCGAGCCGGGTGGGGAAGATCACT 967
Qy 261 AspSerHisGlnTrrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
Db 968 GATTCAACCAAGTGAACACCTTGCTAAAGTTTGATTAACCGCAATTTTATTACTCA 1027
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleYsrThrAla 300
Db 1028 CGCAGCGCAGAGGTGGCCGAGTCGCCACCCCGTTATGGATTTGATCAACAGACGCG 1087
Qy 301 LeuThrProHisProProGlnIlysglnAlaTyrGlyValIleThrLeuProThrSerValLeu 320
```

Db 1088 TTGACGCCCCATCCACCGCAAAAACAGCGGTATGGTGAACATTACCCACTTCAGTGC 1147
 Qy 321 PheilealaglyHisapthrasenleuAlaenleuGlyValaleuGluLeuAantp 340
 Db 1148 TTTATTGCGGACACATACTAATCTGGCAAACTCGCGCGCCACTGAGCTCAACTG 1207
 Qy 341 ThrleuProglYglInProAspAanthrProProglYglYgluLeuValPhegluAgtP 360
 Db 1208 AGCGCTCCAGGTCAAGCCGATACACAGCCGCAAGGTGGAATGTTTGAACCTG 1267
 Qy 361 ArgHAgLeuSerAspAsnSerGlnTppIleGlnValSerleuValPheGlnThrleuGln 380
 Db 1268 CCGCGCTAAGCGATACAGCCAGTGAATTCAGGTTTCGCTGCTTCAGACTTACAG 1327
 Qy 381 GlnMeArgAspLysThrProleuSerleuAanthrProProglYgluValLysleuThr 400
 Db 1328 CAGATGCGATAAACGCCGCTATCATTAATACCGCCGAGAGGTGAACCTGACC 1387
 Qy 401 LeuAlaGlyCyegluGluuArgAsnAlaGlnGlyMetCySerleuAlaGlyPheThrGln 420
 Db 1388 CTGGCAGATGTGAAGAGCAATGCGCAGGCGATGTCTTCTTGCCGCTTTACGCA 1447
 Qy 421 llevalaenGluAlaArgIleProAlaCySserleu 432
 Db 1448 ATCGTAATGAAGCGCGCATACCGGCGTGCAGTTTG 1483

RESULT 14

US-08-910-798-1

Sequence 1, Application US/08910798

Patent No. 5876997

GENERAL INFORMATION:

APPLICANT: KRETTZ

TITLE OF INVENTION: NOVEL PHYTASE

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSER: Fish & Richardson P.C.

CITY: La Jolla

STATE: California

COUNTRY: US

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/910,798

FILING DATE: August 13, 1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: HAILE, PH.D., LISA A.

REGISTRATION NUMBER: 38,347

REFERENCE/DOCKET NUMBER: 09010/029001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619/678-5070

TELEFAX: 619/678-5099

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1272 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

IMMEDIATE SOURCE:

CLONE: PHYTASE

FEATURE:

NAME/KEY:

LOCATION:

US-08-910-798-1

Alignment Scores:

	Pred. No.:	1,31e-210	Length:	1272
	Score:	2063.50	Matches:	405
	Percent Similarity:	93.8%	Mismatches:	0
	Best Local Similarity:	93.8%	Indels:	17
	Query Match:	91.4%	Gaps:	1
	DB:	2		
	US-10-601-319-10 (1-432) x US-08-910-798-1 (1-1272)			
Qy	1	MethylalaleuIleProPheleuSerleuLeuIleProleuThrProGlnSerAla	20	
Db	1	ATGAAGCGATCTTATCCATTTTATCTTCTGATCCGTTAACCCGCAATCTGCA	60	
Qy	21	PheAlaGlnSerGluProGluLeuLeuGlnSerValIleValSerArgIleGly	40	
Db	61	TTGCTGCAAGTGAACCCGAGCTGAAGCTGGAAGATGAGTATGTCAGTGTCAATG	120	
Qy	41	ValAlaGAlaProThrLysAlaIleThrGlnLeuMetGlnAspValIleProAspAlaTrpPro	60	
Db	121	GTGCGTGTCCCAACCAAGGCCAGCACTGATGAGATGTCACCCGACGATGCGCA	180	
Qy	61	ThrTrpProValLysLeuGlyGluLeuThrProArgIlyGlyGluLeuIleAlaTyrlu	80	
Db	181	ACCTGCGCGTAAACTGGGTTGGCTGACACCGCGAGGTGGTGAATCGCTATCTC	240	
Qy	81	GlyHisTyrrTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCyseLysCyPro	100	
Db	241	GGACATTTACCAACGCGACGCTGTGTACCGACGATTTGCTGGCAAAAAGGCTGCCG	300	
Qy	101	GlnSerGlyGlnValAlaIleIleAlaAspValaAspGlnArgThrArgLysThrGlyGlu	120	
Db	301	CAGTGTGTCAGTGTGCGATTTTCTGATGTCCAGCGAGCTTACCCGTTAAACAGGC	357	
Qy	121	AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr	140	
Db	358	-----CAGCGCATACG 369		
Qy	141	SerSerProAspProleuPheAsnProleuLysThrGlyValCyseGlnLeuAspAla	160	
Db	370	TCCAGTCCCGATCCGTTATTTAATCTCTTAATAACCTGGCGTTTGCATGATTAACCG	429	
Qy	161	AsnValThrAspAlaIleLeuGluuArgAlaGlySerIleAlaAspPheThrGlyHis	180	
Db	430	AACGTGATGACGCGATCTTCAAGGACAGGAGGTCAATTGCTGACTTTAACCGGCA	489	
Qy	181	TyrGlnThrAlaPheArgGluLeuGluuArgValLeuAsnPheProGlnSerAsnLysCy	200	
Db	490	CGGCAAACGCGGTTCCGGAACGGAACGGGTGTTATTTCCGCATCAAACTTGTC	549	
Qy	201	LeuLysArgGlyLysGlnAspGlnSerCyseSerleuThrGlnAlaLeuProSerGluLeu	220	
Db	550	CTTAAACGTGAACACAGGAGAAAGCTGTTCATTACCGCAGCATTCATCGGAACTC	609	
Qy	221	LysValSerAlaAspCyseValSerleuThrGlyAlaValSerleuAlaSerMetleuThr	240	
Db	610	AAGGTAGCGCGCAATGTCTTCAATCCAGTGGGTAACTTCGACATCAATGTGACG	669	
Qy	241	GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyIlyArgIleThr	260	
Db	670	GAGATATTCTCCGCAACAAAGCAACAGGAAATGCGGAGCGGGGTGGGAAGATCC	729	
Qy	261	AspSerHisGlnTrpAsnThrLeuLeuSerleuHisAsnAlaGlnPheAspLeuLeuGln	280	
Db	730	GATTCACACCAAGTGAACACCTTCTAAGTTGATTAACGCGCAATTTTATCTCTCA	789	
Qy	281	ArgThrProGluValAlaArgSerArgAlaThrProleuAspLeuIleLysThrAla	300	
Db	790	CGACGCGCAAGGTTGCGCGAGCGCGCACCCCGTTATTTGATGATATGCGACG	849	
Qy	301	LeuThrProHisProProGlnLysGlnAlaTyrlValThrleuProThrSerValleu	320	
Db	850	TTGACGCCCATCCACGCAAAAACAGCGATGATGTGATACCATTCAGTACTG 909		

```
QY 321 PheIleAglYhIaSPThrAsnLeuAlaasnleuGlyValaIeLeuGluLeuAsnTrp 340
Db 910 TTTATGCGGACACGATTAATCTGGCAATCTGGCGGCGACCTGAGACTCAACTGG 969
QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyLeuValaPheGluArgTrp 360
Db 970 ACGCTTCCCGGTGACCGCGATTAACCGCGCGAGGTGTGAACCTGTGTTGAACGCTGG 1029
QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValaPheGlnThrLeuGln 380
Db 1030 CGTGGCTTAACCGATTAACGACGAGTGAATTCAGTTTCGTGCTTTCAGACTTTTACAG 1089
QY 381 GlnMetArgAspLeuThrProLeuSerLeuAsnThrProProGlyGlyValaIysLeuThr 400
Db 1090 CAGATGCGTATTAACCGCCCTGTCTTAATACCGCCCGGAGAGGTGAATCTGACC 1149
QY 401 LeuAlaGlyCysGlyGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1150 CTGGCAGGATGTGAAGACGAAATCGCAGGGCATGTTCGTTGGCAGGTTTACGCAA 1209
QY 421 IleValaAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1210 ATCGTGAATGAAGACGCAATACCGCGCTGCGAGTTTG 1245

RESULT 15
US-09-489-039A-341
; Sequence 341, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709, 2004001
; CURRENT APPLICATION NUMBER: US/09/489, 039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 341
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-341

Alignment Scores:
Pred. No.: 4,93e-48 Length: 1266
Score: 543.00 Matches: 136
Percent Similarity: 49.4% Conservative: 70
Best Local Similarity: 32.6% Mismatches: 181
Query Match: 24.0% Indels: 30
Gaps: 10

US-10-601-319-10 (1-432) x US-09-489-039A-341 (1-1266)
QY 19 SerAlaPheAlaGlnSerGlu-----ProGlu---LeuLysLeuGluSerVal 33
Db 67 TCCGTCGCGCTCAGCGCCGACAGCAAGCCCGCCGAGGGGTATCAGCTCAGAGAGTG 126
QY 34 ValIleValSerArgHisGlyValaArgAlaPro---ThrLysAlaThrGlnLeuMetGln 52
Db 127 CTGATCATGAGCCGCTCAACACTGCGTGGCCGTTGGCCCAATTAACGAGCGAGTGTGAA 186
QY 53 AspValThrProAspAlaThrProThrTrpProValLysLeuGlyGluLeuThrProArg 72
Db 187 CAGTCACCGCCCAAGCCCTGGCCGAGTGGATGTCGGGGGGGCGACCTGACCAACAA 246
QY 73 GlyGlyGluLeuIleAlaIleuGlyHisIleTrpArgGlnArgLeuValaAlaAspGly 92
Db 247 GCGCGTGTGCTGAGAGTGTATATGGCCATATATGCCGAAATGCTGGCCGAGCAAG 306
QY 93 LeuLeuProLysCysGlyCysProGlnSerGlyGlnValaIleAlaIleAlaAspVala 112
Db 307 CTGGTGAACCAAGCGGAGTGGCCGCGGAAACGCGGTTTATGCTTACGCTTAACAGCTG 366
```

```
QY 113 GlnArgThrArgLysThrGlyGluAlaPheAlaAlaGlyLeuAlaProAspCysAlaIle 132
Db 367 CAGCGCACCGTCGCGCACCGGAGCTTCTTATACCGGCGGCTTCCGGGCTGGGCATC 426
QY 133 ThrValHisThrGlnAlaAspThrSerSerProAspProLeuPheAsnProLeuLysThr 152
Db 427 CCGGTGATCAACAGCGCTCAAGATGGGACATGGACCGCACTTCAACCGGTTATTATACC 486
QY 153 GlyValCysGlnLeuAspAsnAlaAsnValThrAspAlaIle---LeuGluArgAlaGly 171
Db 487 GACGACTCCCGCGGCTTCCGAGAAAGCGCTGCAGGCGATGAGAAAGAGCGCCAGGGA 546
QY 172 GlySerIleAlaAspPheThrGlyHisIleTrpGlnThrAlaPheArgGluLeuVal 191
Db 547 ATGCAATTAACCGAG-----AGCTACAAAGCTGCTGAGACGATG 585
QY 192 LeuAsnPheProGlnSerAsnLeuCysLeuLysArgGluLysGlnAspGluSerCysSer 211
Db 586 ATTGACTAACCGCAACTGCGCCCTCTGC-----AAAGAGAAAGGTCTGTTC 633
QY 212 LeuThrGlnAlaLeuProSerGluLeuLysValSerAlaAspCysValSerLeuThrGly 231
Db 634 CTGAGCGAGGTTAAAGATACCTTTAGCCCGGTTATCAGACAGGCGGAGTGTCCGGA 693
QY 232 AlaValSerLeuAlaSerMetLeuThrGluIlePheLeuLeuGlnAlaGlnGlyMet 251
Db 694 CCGTTGAAGTGGGCAACTGCTGTGTAGACCCCTTACCTGTCAATTTATGAAGCTTC 753
QY 252 Pro-----GluProGlyTrpGlyArgIleThrAspSerHisGlnTrpAsnThrLeuLeu 269
Db 754 CCGAAAGATCAGAGTGGGTGGGGGAGATGCCAGCATGAAGTGGCGAGTCTGTGC 813
QY 270 SerLeuHisAsnAlaGlnPheAspLeuLeuGlnArgThrProGluValaIaAspSerArg 289
Db 814 AAGCTGAAATGTGCTACAGATAGTCTGTTAACCTCGGTGGGCGGCAAAAGCTC 873
QY 290 AlaThrProLeuLeuAspLeuIleLysThrAlaLeuThrProHisProProGlnLysGln 309
Db 874 GCCAAACCGCTGTAAATATATGATTAACCGCTGTGGGCGAGGAGCCAGCAAG--- 930
QY 310 AlaTrpGlyValThrLeuProThrSerValLeuPheIleAlaGlyHisAspThrAsnLeu 329
Db 931 -----GCCAAAGTCAACGCTGCTGGTGGGACACGACTCGAATATC 969
QY 330 AlaAsnLeuGlyGlyAlaLeuGluLeuAsn---TrpThrLeuProGlyGlnProAspAsn 348
Db 970 GCTTCGCTGCTCACCGCCCTGAGACTTTAAGCCCTTATCAACTGCGCCGCAGTATAGCGC 1029
QY 349 ThrProProGlyGlyGluLeuValaPheGluArgTrpArgArgLeuSerAspAsnSerGln 368
Db 1030 ACTCTATCGCGGCAAACTGCTTTCAGCGCTGGCATGACAGCGCGGCAATCGCGAT 1089
QY 369 TrpIleGlnValSerLeuValaPheGlnThrLeuGlnGlnMetArgAspLysThrProLeu 388
Db 1090 CTGATGAAGATTTGATGTCTATCAAAATGACCAAGTACGCTTAACGCGGAGCGCTTA 1149
QY 389 SerLeuAsnThrProProGlyGlyValaLysLeuThrLeuAlaGlyCysGlyGluArgAsn 408
Db 1150 ACCCTGAGGCGCCGCGCAGCGGCTCACTTGCCTTAACGCGCTGTCCGTT--GAT 1206
QY 409 AlaGlnGlyMetCysSerLeuAlaGlyPheThrGlnIleValaAsnGluAla 425
Db 1207 GACCAAGGATTTGCGCGCTGGAACCTTTAAATGTGATCAACAGAGCG 1257
```

Search completed: June 14, 2006, 15:34:57
Job time : 248.459 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2006, 10:45:37 ; Search time 140.697 Seconds
(without alignments)
1422.265 Million cell updates/sec

Title: US-10-601-319-10
Perfect score: 2258
Sequence: 1 MKAILIPFLSLILPLTPQSA.....CSLAGFTQVNEARIPACSL 432

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:
1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2258	100.0	432	4	US-10-601-319-10 Sequence 10, Appl
2	2258	100.0	432	5	US-10-933-115-10 Sequence 10, Appl
3	2258	100.0	432	6	US-11-056-354-2 Sequence 2, Appl
4	2258	100.0	432	4	US-10-156-660-2 Sequence 2, Appl
5	2258	98.7	430	3	US-09-866-379-10 Sequence 10, Appl
6	2182	96.6	432	3	US-09-866-379-8 Sequence 8, Appl
7	2182	96.6	432	4	US-10-156-660-4 Sequence 4, Appl
8	2182	96.6	432	4	US-10-282-122A-43351 Sequence 4, Appl
9	2182	96.6	432	4	US-10-601-319-8 Sequence 4, Appl
10	2182	96.6	432	4	US-10-472-317-41 Sequence 4, Appl
11	2182	96.6	432	5	US-10-933-115-8 Sequence 8, Appl
12	2182	96.6	432	6	US-11-056-354-4 Sequence 4, Appl
13	2177	96.4	432	6	US-11-018-709-1 Sequence 2, Appl
14	2171	96.1	440	3	US-09-777-566A-2 Sequence 2, Appl
15	2171	96.1	440	3	US-09-866-379-2 Sequence 2, Appl
16	2171	96.1	440	4	US-10-434-985-2 Sequence 2, Appl
17	2171	96.1	440	4	US-10-430-356-2 Sequence 2, Appl
18	2171	96.1	440	4	US-10-601-319-2 Sequence 2, Appl
19	2171	96.1	440	5	US-10-933-115-2 Sequence 12, Appl
20	2168.5	96.0	437	6	US-11-074-522-12 Sequence 12, Appl
21	2157	95.5	412	4	US-10-334-672-1 Sequence 1, Appl
22	2157	95.5	412	4	US-10-334-672-1 Sequence 1, Appl
23	2153	95.5	432	6	US-11-018-709-3 Sequence 3, Appl
24	2153	95.3	410	4	US-10-021-723A-13 Sequence 13, Appl
25	2147	95.1	432	4	US-10-284-962-3 Sequence 3, Appl
26	2147	95.1	432	4	US-10-284-962-3 Sequence 3, Appl
27	2147	95.1	432	4	US-10-284-962-14 Sequence 14, Appl

28	2147	95.1	433	4	US-10-266-041-1 Sequence 1, Appl
29	2144	95.0	432	4	US-10-284-962-5 Sequence 5, Appl
30	2112	93.5	432	4	US-10-021-723A-16 Sequence 16, Appl
31	1192	52.8	261	5	US-10-450-763-56057 Sequence 56057, A
32	976.5	43.2	441	4	US-10-021-723A-12 Sequence 12, Appl
33	959.5	42.5	441	4	US-10-282-122A-77792 Sequence 77792, A
34	951.5	42.1	440	4	US-10-021-723A-2 Sequence 2, Appl
35	948.5	42.0	420	4	US-10-021-723A-4 Sequence 4, Appl
36	782	34.6	146	4	US-10-021-723A-10 Sequence 10, Appl
37	637	28.2	144	5	US-10-450-763-54615 Sequence 54615, A
38	627.5	27.8	318	4	US-10-021-723A-14 Sequence 14, Appl
39	475.5	21.1	409	4	US-10-021-723A-8 Sequence 8, Appl
40	462.5	20.5	421	4	US-10-021-723A-6 Sequence 6, Appl
41	414	18.3	441	6	US-11-098-686-10683 Sequence 10683, A
42	310	13.7	118	5	US-10-450-763-56055 Sequence 56055, A
43	197	8.7	426	4	US-10-257-174-44 Sequence 44, Appl
44	197	8.7	426	4	US-10-343-357-6 Sequence 6, Appl
45	197	8.7	426	5	US-10-450-763-53707 Sequence 53707, A

ALIGNMENTS

```

RESULT 1
US-10-601-319-10
; Sequence 10, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
; FILE REFERENCE: 09010-029011
; CURRENT APPLICATION NUMBER: US/10/601,319
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: modified phyase enzyme
US-10-601-319-10
Query Match 100.0%; Score 2258; DB 4; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.86-196;
Matches 432; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKAILIPFLSLILPLTPQSAFAQSEPELKLTSVTVSHGVRAPTKATQLMQDVPDMP 60
DB 1 MKAILIPFLSLILPLTPQSAFAQSEPELKLTSVTVSHGVRAPTKATQLMQDVPDMP 60
QY 61 TWPVLDGELTPRGGLIAYLVGHYWRQRLVADGLPKCCGPGCGVAILADVDERTKTCGE 120
DB 61 TWPVLDGELTPRGGLIAYLVGHYWRQRLVADGLPKCCGPGCGVAILADVDERTKTCGE 120

```

Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTVGCQOLDNANVTDAIIRAGGSIAIDFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTVGCQOLDNANVTDAIIRAGGSIAIDFTGH 180
Qy 181 YQTAFARELERVLTNPQSNLCLKREKQDSCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
Db 181 YQTAFARELERVLTNPQSNLCLKREKQDSCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
Qy 241 EIFLLQQAQGMPEBPGWGRITDISHQWNTLLSHNAQFDLLQRTPEVARSRAATPLDLIXTA 300
Db 241 EIFLLQQAQGMPEBPGWGRITDISHQWNTLLSHNAQFDLLQRTPEVARSRAATPLDLIXTA 300
Qy 301 LTPHPQKQAYGVTLPTSVLPIAGHDTNLAMIGALBLNMTLLPGQPDNTPPGSELVPERW 360
Db 301 LTPHPQKQAYGVTLPTSVLPIAGHDTNLAMIGALBLNMTLLPGQPDNTPPGSELVPERW 360
Qy 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPPGEVYKLLTAGCBERNAQGMCSLAGFTQ 420
Db 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPPGEVYKLLTAGCBERNAQGMCSLAGFTQ 420
Qy 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432

RESULT 2
US-10-933-115-10

; Sequence 10, Application US/10933115
; Publication No. US20050281792A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O'Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES
; TITLE OF INVENTION: THEREOP
; FILE REFERENCE: 564462001822
; CURRENT APPLICATION NUMBER: US/10/933,115
; CURRENT FILING DATE: 2004-09-01
; PRIOR APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: modified phytase enzyme
US-10-933-115-10

Query Match 100.0%; Score 2258; DB 5; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.8e-196;
Matches 432; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKAIIPLSLILITLTPQSAFQSEPELKLBSVIVSRHGVRAPTKATQMODVTTPDAMP 60
Db 1 KKAIIPLSLILITLTPQSAFQSEPELKLBSVIVSRHGVRAPTKATQMODVTTPDAMP 60

Qy 61 TWPVKGLSPRGSELIAVLGHYWRORLVADGLPKCGCPQSGQVAIIVADBERTRTKGE 120
Db 61 TWPVKGLSPRGSELIAVLGHYWRORLVADGLPKCGCPQSGQVAIIVADBERTRTKGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTVGCQOLDNANVTDAIIRAGGSIAIDFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTVGCQOLDNANVTDAIIRAGGSIAIDFTGH 180
Qy 181 YQTAFARELERVLTNPQSNLCLKREKQDSCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
Db 181 YQTAFARELERVLTNPQSNLCLKREKQDSCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
Qy 241 EIFLLQQAQGMPEBPGWGRITDISHQWNTLLSHNAQFDLLQRTPEVARSRAATPLDLIXTA 300
Db 241 EIFLLQQAQGMPEBPGWGRITDISHQWNTLLSHNAQFDLLQRTPEVARSRAATPLDLIXTA 300
Qy 301 LTPHPQKQAYGVTLPTSVLPIAGHDTNLAMIGALBLNMTLLPGQPDNTPPGSELVPERW 360
Db 301 LTPHPQKQAYGVTLPTSVLPIAGHDTNLAMIGALBLNMTLLPGQPDNTPPGSELVPERW 360
Qy 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPPGEVYKLLTAGCBERNAQGMCSLAGFTQ 420
Db 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPPGEVYKLLTAGCBERNAQGMCSLAGFTQ 420
Qy 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432

RESULT 3
US-11-056-354-2

; Sequence 2, Application US/11056354
; Publication No. US20050246780A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson R.
; APPLICANT: Garrett, James B.
; APPLICANT: O'DONOGHUE, Eileen
; APPLICANT: MATHUR, ERIC J.
; TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM AND METHODS
; TITLE OF INVENTION: FOR MAKING AND USING THEM
; FILE REFERENCE: 564462001803/D1370-9C1
; CURRENT APPLICATION NUMBER: US/11/056,354
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 10/156,660
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Modified phytase enzyme
US-11-056-354-2

Query Match 100.0%; Score 2258; DB 6; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.8e-196;
Matches 432; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAAILIPFLSLILPLTPQSAFAQSBPELKLSEVIVYSRHGVRAPTKATOLMODVTPDAMP	60
Db	1	MAAILIPFLSLILPLTPQSAFAQSBPELKLSEVIVYSRHGVRAPTKATOLMODVTPDAMP	60
Qy	61	TWPKLGEILTPRGSGELIAYLGHYWRQRLVADGLLPKCGCPQSQGVAILADVDERTRKTGE	120
Db	61	TWPKLGEILTPRGSGELIAYLGHYWRQRLVADGLLPKCGCPQSQGVAILADVDERTRKTGE	120
Qy	121	APAAGLAAPCCATVTWHQADPTSSDPLPNPLKTYGCVGLDNNATVDALIERAGSIADPTGH	180
Db	121	APAAGLAAPCCATVTWHQADPTSSDPLPNPLKTYGCVGLDNNATVDALIERAGSIADPTGH	180
Qy	181	YOTAFELERVLNFPQSNCLKKEKODESCSLTOALPSELKVSADCVSLTGAVSIASMLT	240
Db	181	YOTAFELERVLNFPQSNCLKKEKODESCSLTOALPSELKVSADCVSLTGAVSIASMLT	240
Qy	241	EIFLLQQAQGMPEPGMGRIITDSHQWNTLLSLHNAQFDLLQRTBEVARSRAPTPLDLIKTA	300
Db	241	EIFLLQQAQGMPEPGMGRIITDSHQWNTLLSLHNAQFDLLQRTBEVARSRAPTPLDLIKTA	300
Qy	301	LTPHPPOKAYGVTLPTSLVFLIGHNTNLANLGALFLAWNTLPGQDPNTPPGSELVFERW	360
Db	301	LTPHPPOKAYGVTLPTSLVFLIGHNTNLANLGALFLAWNTLPGQDPNTPPGSELVFERW	360
Qy	361	RLSDNSQWIOVSLVYFQTLQOMKDKTPLSLNTPPGEVKULLACGBERNAGMGLSAGFTQ	420
Db	361	RLSDNSQWIOVSLVYFQTLQOMKDKTPLSLNTPPGEVKULLACGBERNAGMGLSAGFTQ	420
Qy	421	IVNEARIPACSL 432	
Db	421	IVNEARIPACSL 432	
RESULT 4			
US-10-156-660-2			
; Sequence 2, Application US/10156660			
; Publication No. US20030103958A1			
; GENERAL INFORMATION:			
; APPLICANT: Short, Jay M.			
; APPLICANT: Kretz, Keith			
; APPLICANT: Gray, Kevin A.			
; APPLICANT: Barton, Nelson R.			
; APPLICANT: Garrett, James B.			
; APPLICANT: O'Donoghue, Eileen			
; APPLICANT: Mathur, Eric J.			
; TITLE OF INVENTION: PHYLASES, NUCLEIC ACIDS ENCODING THEM			
; FILE REFERENCE: 09010-029007			
; CURRENT APPLICATION NUMBER: US/10/156,660			
; CURRENT FILING DATE: 2002-10-01			
; PRIOR APPLICATION NUMBER: US 09/866,379			
; PRIOR FILING DATE: 2001-05-24			
; PRIOR APPLICATION NUMBER: US 09/580,515			
; PRIOR FILING DATE: 2000-05-25			
; PRIOR APPLICATION NUMBER: US 09/318,528			
; PRIOR FILING DATE: 1999-05-25			
; PRIOR APPLICATION NUMBER: US 09/291,931			
; PRIOR FILING DATE: 1999-04-13			
; PRIOR APPLICATION NUMBER: US 09/259,214			
; PRIOR FILING DATE: 1999-03-01			
; PRIOR APPLICATION NUMBER: US 08/910,798			
; PRIOR FILING DATE: 1997-08-13			
; NUMBER OF SEQ ID NOS: 4			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 2			
; LENGTH: 436			
; TYPE: PRT			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: modified phytase enzyme			
US-10-156-660-2			

```

Query Match          100.0%; Score 2258; DB 4; Length 436;
Best Local Similarity 100.0%; Pkcd No. 1,8e-196;
Matches 432; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MKAILPPLSLIPLTPQSAFAQSEBPEIKLESVIVSRHGVRAPTKATQLMQDVTDPAMP 60
DB      1  MKAILPPLSLIPLTPQSAFAQSEBPEIKLESVIVSRHGVRAPTKATQLMQDVTDPAMP 60

QY      61  TWPVYDGLSTRGGELIAYLGHYRQRILVADGILLPKGCCPQSGQVAILADVDETRKTGE 120
DB      61  TWPVYDGLSTRGGELIAYLGHYRQRILVADGILLPKGCCPQSGQVAILADVDETRKTGE 120

QY      121  AFAAGLAPDCAITYHTQADTSSPPPLFNPPLKTYGCOLDNAVVTALIERAGGSYADFTGH 180
DB      121  AFAAGLAPDCAITYHTQADTSSPPPLFNPPLKTYGCOLDNAVVTALIERAGGSYADFTGH 180

QY      181  YQTAFREBLERVLPQSNLCLKREKQDECSLTQALPSELKVASADCVSLTGAVSLASMLT 240
DB      181  YQTAFREBLERVLPQSNLCLKREKQDECSLTQALPSELKVASADCVSLTGAVSLASMLT 240

QY      241  EIFLLQQAQGMPEBQWGRITDSHQMNTLLSLHNAQFDLLQRTPEVANSRATPLDLITXA 300
DB      241  EIFLLQQAQGMPEBQWGRITDSHQMNTLLSLHNAQFDLLQRTPEVANSRATPLDLITXA 300

QY      301  LTPHPPOKQAGVTLPTSVLFIAGHDNTLAMIAGALBELNMTLPGOPDPTPGGELVPERW 360
DB      301  LTPHPPOKQAGVTLPTSVLFIAGHDNTLAMIAGALBELNMTLPGOPDPTPGGELVPERW 360

QY      361  RRLSDNSQWIVSVLFQTLQCMRDKTPLSLNTPGGEVYLTLAGEBERNAQCMCSLAGFTQ 420
DB      361  RRLSDNSQWIVSVLFQTLQCMRDKTPLSLNTPGGEVYLTLAGEBERNAQCMCSLAGFTQ 420

QY      421  IYNEARIPACSL 432
DB      421  IYNEARIPACSL 432

RESULT 5
US-09-866-379-10
/ Sequence 10, Application US/09866379
/ Patent No. US20020136754A1
/ GENERAL INFORMATION:
/ APPLICANT: DIVERSA CORPORATION
/ APPLICANT: SHORT, Jay
/ APPLICANT: KRETZ, Keith
/ APPLICANT: GRAY, Kevin
/ APPLICANT: BARTON, Nelson
/ APPLICANT: GARRETT, James
/ APPLICANT: O'DONOGHIE, Bileen
/ TITLE OR INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
/ FILE REFERENCE: DIVER1370-7
/ CURRENT APPLICATION NUMBER: US/09/866,379
/ PRIOR FILING DATE: 2001-05-24
/ PRIOR APPLICATION NUMBER: US 09/580,515
/ PRIOR FILING DATE: 2000-05-25
/ PRIOR APPLICATION NUMBER: US 09/318,528
/ PRIOR FILING DATE: 1999-05-25
/ PRIOR APPLICATION NUMBER: US 09/291,931
/ PRIOR FILING DATE: 1999-04-13
/ PRIOR APPLICATION NUMBER: US 09/259,214
/ PRIOR FILING DATE: 1999-03-01
/ PRIOR APPLICATION NUMBER: US 08/910,798
/ PRIOR FILING DATE: 1997-08-13
/ NUMBER OF SEQ ID NOS: 10
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 10
/ LENGTH: 430
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Modified phytase
US-09-866-379-10

```

Query Match 98.7%; Score 2228; DB 3; Length 430;
Best Local Similarity 99.5%; Pred. No. 9.7e-194;
Matches 430; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

```
QY 1 MKAILPPLSLILPLTPQSAFAQSEPELKESVIVSRHGVRATPKATQTMQDVTPDAMP 60
DB 1 MKAILPPLSLILPLTPQSAFAQSEPELKESVIVSRHGVRATPKATQTMQDVTPDAMP 60
QY 61 TWPVKGEIPLPRGELLAYIGHYWRQRLVADGLPKCGCPQSGQVAIIADVDETRKTGE 120
DB 61 TWPVKGEIPLPRGELLAYIGHYWRQRLVADGLPKCGCPQSGQVAIIADVDETRKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTVGCQDNNANVTDAIIERAGGSIAADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTVGCQDNNANVTDAIIERAGGSIAADFTGH 179
QY 181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVALSASMLT 240
DB 180 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVALSASMLT 239
QY 241 EIFLLQQAQGMPEBGMGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRAATPLDLIKTA 300
DB 240 EIFLLQQAQGMPEBGMGRITDSHQWNTLLSLHNAQFD-LQRTPEVARSRAATPLDLIKTA 298
QY 301 LTPHPQKQAYGVTLPTSVLFLIAGHDYTLNMLGALBLNMTLPGQPDNTPPGGELVFERW 360
DB 299 LTPHPQKQAYGVTLPTSVLFLIAGHDYTLNMLGALBLNMTLPGQPDNTPPGGELVFERW 358
QY 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPGGEVYKTLTAGCERNNAQMCSLAGFTQ 420
DB 359 RRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPGGEVYKTLTAGCERNNAQMCSLAGFTQ 418
QY 421 IVNEARIPACSL 432
DB 419 IVNEARIPACSL 430
```

RESULT 6
US-09-866-379-8
Sequence 8, Application US/09866379
Patent No. US20020136754A1

GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: SHORT, Jay
APPLICANT: KRETZ, Keith
APPLICANT: GRAY, Kevin
APPLICANT: NELSON, Barton
APPLICANT: GARRETT, James
APPLICANT: O'DONOGHUE, Bileen
TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
FILE REFERENCE: DIVER1370-7
CURRENT APPLICATION NUMBER: US/09/866,379
PRIOR FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: US 09/580,515
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 09/318,528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: US 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 09/259,214
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: US 08/910,798
PRIOR FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.1
SEQ ID NO 8
LENGTH: 432
TYPE: PRT
ORGANISM: Escherichia coli
US-09-866-379-8

Query Match 96.6%; Score 2182; DB 3; Length 432;
Best Local Similarity 98.1%; Pred. No. 1.5e-189;

Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
QY 1 MKAILPPLSLILPLTPQSAFAQSEPELKESVIVSRHGVRATPKATQTMQDVTPDAMP 60
DB 1 MKAILPPLSLILPLTPQSAFAQSEPELKESVIVSRHGVRATPKATQTMQDVTPDAMP 60
QY 61 TWPVKGEIPLPRGELLAYIGHYWRQRLVADGLPKCGCPQSGQVAIIADVDETRKTGE 120
DB 61 TWPVKGEIPLPRGELLAYIGHYWRQRLVADGLPKCGCPQSGQVAIIADVDETRKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTVGCQDNNANVTDAIIERAGGSIAADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTVGCQDNNANVTDAIIERAGGSIAADFTGH 180
QY 181 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVALSASMLT 240
DB 180 YQTAFRELERVLNFPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVALSASMLT 240
QY 241 EIFLLQQAQGMPEBGMGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRAATPLDLIKTA 300
DB 241 EIFLLQQAQGMPEBGMGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRAATPLDLIKTA 300
QY 301 LTPHPQKQAYGVTLPTSVLFLIAGHDYTLNMLGALBLNMTLPGQPDNTPPGGELVFERW 360
DB 301 LTPHPQKQAYGVTLPTSVLFLIAGHDYTLNMLGALBLNMTLPGQPDNTPPGGELVFERW 360
QY 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPGGEVYKTLTAGCERNNAQMCSLAGFTQ 420
DB 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPGGEVYKTLTAGCERNNAQMCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432
```

RESULT 7
US-10-156-660-4
Sequence 4, Application US/10156660
Publication No. US20030103958A1

GENERAL INFORMATION:
APPLICANT: Short, Jay M.
APPLICANT: Kretz, Keith
APPLICANT: Gray, Kevin A.
APPLICANT: Barton, Nelson R.
APPLICANT: Garrett, James B.
APPLICANT: O'Donoghue, Bileen
TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM
FILE REFERENCE: 09010-029007
CURRENT APPLICATION NUMBER: US/10/156,660
PRIOR FILING DATE: 2002-10-01
PRIOR APPLICATION NUMBER: US 09/866,379
PRIOR FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: US 09/580,515
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 09/318,528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: US 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 09/259,214
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: US 08/910,798
PRIOR FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows version 4.0
SEQ ID NO 4
LENGTH: 432
TYPE: PRT
ORGANISM: Escherichia coli
US-10-156-660-4

Query Match 96.6%; Score 2182; DB 4; Length 432;

Best Local Similarity 98.1%; Pred. No. 1.5e-189;
Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 MRAIIIPFSLIPLTPQSAFAQSEBELKESVIVSRHGVARPTATQTMQDVTDPDAMP 60
DB 1 MRAIIIPFSLIPLTPQSAFAQSEBELKESVIVSRHGVARPTATQTMQDVTDPDAMP 60

QY 61 TWPVKLGELTPRGSELIAVIGHYWRQRLVADGLLPKCGCPQSGQVAIADVDERTKKTGE 120
DB 61 TWPVKLGELTPRGSELIAVIGHYWRQRLVADGLLPKCGCPQSGQVAIADVDERTKKTGE 120

QY 121 APAAGIAPPCAITVHTQADTSSPDLFNPPLKTVGCQLDANVTDAILERAGGSIAIDFTGH 180
DB 121 APAAGIAPPCAITVHTQADTSSPDLFNPPLKTVGCQLDANVTDAILERAGGSIAIDFTGH 180

QY 181 YGTAFRELERVUNFPQSNICLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 YGTAFRELERVUNFPQSNICLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

QY 241 EIFLLQQAQMPGPGMGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIXTA 300
DB 241 EIFLLQQAQMPGPGMGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIXTA 300

QY 301 LTPHPQKQAYGVTLPSTVLFIAHDTNLANIGALELWNTLPGQPDNTPPGSELVFERM 360
DB 301 LTPHPQKQAYGVTLPSTVLFIAHDTNLANIGALELWNTLPGQPDNTPPGSELVFERM 360

QY 361 RLSDNSQWIOVSLVFQTLQQRDKTPLUSLNTPPGSEVKLTLAGCERNNAQMCSLAGFTQ 420
DB 361 RLSDNSQWIOVSLVFQTLQQRDKTPLUSLNTPPGSEVKLTLAGCERNNAQMCSLAGFTQ 420

QY 421 IYNEARIPACSL 432
DB 421 IYNEARIPACSL 432

RESULT 8
US-10-282-122A-43351
; Sequence 43351, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangau
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Foreyth, R.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43351
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-282-122A-43351

Query Match
Best Local Similarity 96.6%; Score 2182; DB 4; Length 432;
Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 MRAIIIPFSLIPLTPQSAFAQSEBELKESVIVSRHGVARPTATQTMQDVTDPDAMP 60
DB 1 MRAIIIPFSLIPLTPQSAFAQSEBELKESVIVSRHGVARPTATQTMQDVTDPDAMP 60

QY 61 TWPVKLGELTPRGSELIAVIGHYWRQRLVADGLLPKCGCPQSGQVAIADVDERTKKTGE 120
DB 61 TWPVKLGELTPRGSELIAVIGHYWRQRLVADGLLPKCGCPQSGQVAIADVDERTKKTGE 120

QY 121 APAAGIAPPCAITVHTQADTSSPDLFNPPLKTVGCQLDANVTDAILERAGGSIAIDFTGH 180
DB 121 APAAGIAPPCAITVHTQADTSSPDLFNPPLKTVGCQLDANVTDAILERAGGSIAIDFTGH 180

QY 181 YGTAFRELERVUNFPQSNICLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 YGTAFRELERVUNFPQSNICLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

QY 241 EIFLLQQAQMPGPGMGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIXTA 300
DB 241 EIFLLQQAQMPGPGMGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIXTA 300

QY 301 LTPHPQKQAYGVTLPSTVLFIAHDTNLANIGALELWNTLPGQPDNTPPGSELVFERM 360
DB 301 LTPHPQKQAYGVTLPSTVLFIAHDTNLANIGALELWNTLPGQPDNTPPGSELVFERM 360

QY 361 RLSDNSQWIOVSLVFQTLQQRDKTPLUSLNTPPGSEVKLTLAGCERNNAQMCSLAGFTQ 420
DB 361 RLSDNSQWIOVSLVFQTLQQRDKTPLUSLNTPPGSEVKLTLAGCERNNAQMCSLAGFTQ 420

QY 421 IYNEARIPACSL 432
DB 421 IYNEARIPACSL 432

RESULT 9
US-10-601-319-8
; Sequence 8, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O'Donoghue, Eileen
; TITLE OF INVENTION: RECOMBINANT PHRYASES AND METHODS OF MAKING
; FILE REFERENCE: 09010-029011
; CURRENT APPLICATION NUMBER: US/10/601,319
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25

;; PRIOR APPLICATION NUMBER: US 09/291,931
;; PRIOR FILING DATE: 1999-04-13
;; PRIOR APPLICATION NUMBER: US 09/259,214
;; PRIOR FILING DATE: 1999-03-01
;; PRIOR APPLICATION NUMBER: US 08/910,798
;; PRIOR FILING DATE: 1997-08-13
;; NUMBER OF SEQ ID NOS: 10
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 8
;; LENGTH: 432
;; TYPE: PRT
;; ORGANISM: Escherichia coli
US-10-601-319-8

Query Match 96.6%; Score 2182; DB 4; Length 432;
Best Local Similarity 98.1%; Pred. No. 1.5e-189;
Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 MKAILIPFLSLILPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQLMQDVTDPAMP 60
DB 1 MKAILIPFLSLILPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQLMQDVTDPAMP 60
QY 61 TWPVGLGELTPRGSELIAVYGHYWRQRLVADGLPKCGCPQSGQVAIIADVDETRKTGE 120
DB 61 TWPVGLGELTPRGSELIAVYGHYWRQRLVADGLPKCGCPQSGQVAIIADVDETRKTGE 120
QY 121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQLDNNAVTDAILERAGGSIADEFTH 180
DB 121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQLDNNAVTDAILERAGGSIADEFTH 180
QY 181 YQTAFRELERVLPQSNLCLKREKODESCSLTQALPSELKVSADCVSLTGAVALASMLT 240
DB 181 YQTAFRELERVLPQSNLCLKREKODESCSLTQALPSELKVSADCVSLTGAVALASMLT 240
QY 241 EIFLLQQAQGMPEBGMGRITDSHQWNTLLSIHNAQFDLQRTPEVARSRATPLDLIKTA 300
DB 241 EIFLLQQAQGMPEBGMGRITDSHQWNTLLSIHNAQFDLQRTPEVARSRATPLDLIKTA 300
QY 301 LTPHPQKQAYGVTLPFSVLFIAGHDTNLAMGALBELNMTLPQCPDNTPPGSELVFERW 360
DB 301 LTPHPQKQAYGVTLPFSVLFIAGHDTNLAMGALBELNMTLPQCPDNTPPGSELVFERW 360
QY 361 RRLSDNSQMIQVSLVFQTLQGMKDKTPLSLNTPPGEVLTLAGCEERNAQMCSLAGFTQ 420
DB 361 RRLSDNSQMIQVSLVFQTLQGMKDKTPLSLNTPPGEVLTLAGCEERNAQMCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 10
US-10-472-317-41
;; Sequence 41, Application US/10472317
;; Publication No. US20040185562A1
;; GENERAL INFORMATION:
;; APPLICANT: Cargill, Incorporated
;; TITLE OF INVENTION: Myo-Inositol Oxygenases
;; FILE REFERENCE: 10829/003US1
;; CURRENT APPLICATION NUMBER: US/10/472,317
;; CURRENT FILING DATE: 2003-09-17
;; PRIOR APPLICATION NUMBER: PCT/US02/08404
;; PRIOR FILING DATE: 2002-03-19
;; PRIOR APPLICATION NUMBER: US 60/277,148
;; PRIOR FILING DATE: 2001-03-19
;; NUMBER OF SEQ ID NOS: 72
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 41
;; LENGTH: 432
;; TYPE: PRT
;; ORGANISM: Escherichia coli
US-10-472-317-41

Query Match 96.6%; Score 2182; DB 4; Length 432;
Best Local Similarity 98.1%; Pred. No. 1.5e-189;
Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 MKAILIPFLSLILPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQLMQDVTDPAMP 60
DB 1 MKAILIPFLSLILPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQLMQDVTDPAMP 60
QY 61 TWPVGLGELTPRGSELIAVYGHYWRQRLVADGLPKCGCPQSGQVAIIADVDETRKTGE 120
DB 61 TWPVGLGELTPRGSELIAVYGHYWRQRLVADGLPKCGCPQSGQVAIIADVDETRKTGE 120
QY 121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQLDNNAVTDAILERAGGSIADEFTH 180
DB 121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQLDNNAVTDAILERAGGSIADEFTH 180
QY 181 YQTAFRELERVLPQSNLCLKREKODESCSLTQALPSELKVSADCVSLTGAVALASMLT 240
DB 181 YQTAFRELERVLPQSNLCLKREKODESCSLTQALPSELKVSADCVSLTGAVALASMLT 240
QY 241 EIFLLQQAQGMPEBGMGRITDSHQWNTLLSIHNAQFDLQRTPEVARSRATPLDLIKTA 300
DB 241 EIFLLQQAQGMPEBGMGRITDSHQWNTLLSIHNAQFDLQRTPEVARSRATPLDLIKTA 300
QY 301 LTPHPQKQAYGVTLPFSVLFIAGHDTNLAMGALBELNMTLPQCPDNTPPGSELVFERW 360
DB 301 LTPHPQKQAYGVTLPFSVLFIAGHDTNLAMGALBELNMTLPQCPDNTPPGSELVFERW 360
QY 361 RRLSDNSQMIQVSLVFQTLQGMKDKTPLSLNTPPGEVLTLAGCEERNAQMCSLAGFTQ 420
DB 361 RRLSDNSQMIQVSLVFQTLQGMKDKTPLSLNTPPGEVLTLAGCEERNAQMCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 11
US-10-933-115-8
;; Sequence 8, Application US/10933115
;; Publication No. US20050281792A1
;; GENERAL INFORMATION:
;; APPLICANT: Short, Jay M.
;; APPLICANT: Kretz, Keith A.
;; APPLICANT: Gray, Kevin A.
;; APPLICANT: Barton, Nelson Robert
;; APPLICANT: Garrett, James B.
;; APPLICANT: O' Donoghue, Eileen
;; APPLICANT: Mathur, Eric J.
;; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES
;; FILE REFERENCE: 56446201822
;; CURRENT APPLICATION NUMBER: US/10/933,115
;; CURRENT FILING DATE: 2004-09-01
;; PRIOR APPLICATION NUMBER: US/09/866,379
;; PRIOR FILING DATE: 2001-05-24
;; PRIOR APPLICATION NUMBER: US 09/580,515
;; PRIOR FILING DATE: 2000-05-25
;; PRIOR APPLICATION NUMBER: US 09/318,528
;; PRIOR FILING DATE: 1999-05-25
;; PRIOR APPLICATION NUMBER: US 09/291,931
;; PRIOR FILING DATE: 1999-04-13
;; PRIOR APPLICATION NUMBER: US 09/259,214
;; PRIOR FILING DATE: 1999-03-01
;; PRIOR APPLICATION NUMBER: US 08/910,798
;; PRIOR FILING DATE: 1997-08-13
;; NUMBER OF SEQ ID NOS: 10
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 8
;; LENGTH: 432
;; TYPE: PRT
;; ORGANISM: Escherichia coli
US-10-933-115-8

```
Query Match      96.6%; Score 2182; DB 5; Length 432;
Best Local Similarity 98.1%; Pred. No. 1.5e-189;
Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY      1 MKAIIIPFLSLIPLTPQSAFAQSEBELKESVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
      1 MKAIIIPFLSLIPLTPQSAFAQSEBELKESVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
Db      61 TWPVKLGELTPRGSELIAVIGHYWRQRLVADGLPKCGCPQSGQVAIADVDETRTKTGE 120
QY      61 TWPVKLGELTPRGSELIAVIGHYWRQRLVADGLPKCGCPQSGQVAIADVDETRTKTGE 120
      61 TWPVKLGELTPRGSELIAVIGHYWRQRLVADGLPKCGCPQSGQVAIADVDETRTKTGE 120
Db      121 APAAGIAPPCATVTHQADTSSPDPLFNPPLKTVGCOLDNANVTDAILBRAGSIA DFTGH 180
QY      121 APAAGIAPPCATVTHQADTSSPDPLFNPPLKTVGCOLDNANVTDAILBRAGSIA DFTGH 180
      121 APAAGIAPPCATVTHQADTSSPDPLFNPPLKTVGCOLDNANVTDAILBRAGSIA DFTGH 180
Db      181 YQTAPELELRLVNFPOSNCLKREKODESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
QY      181 YQTAPELELRLVNFPOSNCLKREKODESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
      181 YQTAPELELRLVNFPOSNCLKREKODESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
Db      241 EIFLLQAOAGMPGPGWGRITDISHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
QY      241 EIFLLQAOAGMPGPGWGRITDISHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
      241 EIFLLQAOAGMPGPGWGRITDISHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
Db      301 LTPHPPOKAYGVTLPTSVLFTAGHDNTLANLGALIELWNTLPGQPDNTPPGSELVFERM 360
QY      301 LTPHPPOKAYGVTLPTSVLFTAGHDNTLANLGALIELWNTLPGQPDNTPPGSELVFERM 360
      301 LTPHPPOKAYGVTLPTSVLFTAGHDNTLANLGALIELWNTLPGQPDNTPPGSELVFERM 360
Db      361 RLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGSEVKTLTAGCERNRQMGCSLAGFTQ 420
QY      361 RLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGSEVKTLTAGCERNRQMGCSLAGFTQ 420
      361 RLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGSEVKTLTAGCERNRQMGCSLAGFTQ 420
Db      421 IYNEARIPACSL 432
QY      421 IYNEARIPACSL 432
      421 IYNEARIPACSL 432
Db      421 IYNEARIPACSL 432

RESULT 12
US-11-056-354-4
; Sequence 4, Application US/11056354
; Publication No. US20050246780A1
; GENERAL INFORMATION:
; APPLICANT: SHORT, Jay M.
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin A.
; APPLICANT: BARTON, Nelson R.
; APPLICANT: GARRETT, James B.
; APPLICANT: O'DONOGHUE, Eileen
; APPLICANT: MATHUR, Eric J.
; TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM AND METHODS
; FILE REFERENCE: 56462001803/D1370-9C1
; CURRENT APPLICATION NUMBER: US/11/056,354
; PRIOR FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 10/156,660
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 432
```

```
; TYPE: prt
; ORGANISM: Escherichia coli appa phytase
; US-11-056-354-4

Query Match      96.6%; Score 2182; DB 6; Length 432;
Best Local Similarity 98.1%; Pred. No. 1.5e-189;
Matches 424; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY      1 MKAIIIPFLSLIPLTPQSAFAQSEBELKESVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
      1 MKAIIIPFLSLIPLTPQSAFAQSEBELKESVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
Db      61 TWPVKLGELTPRGSELIAVIGHYWRQRLVADGLPKCGCPQSGQVAIADVDETRTKTGE 120
QY      61 TWPVKLGELTPRGSELIAVIGHYWRQRLVADGLPKCGCPQSGQVAIADVDETRTKTGE 120
      61 TWPVKLGELTPRGSELIAVIGHYWRQRLVADGLPKCGCPQSGQVAIADVDETRTKTGE 120
Db      121 APAAGIAPPCATVTHQADTSSPDPLFNPPLKTVGCOLDNANVTDAILBRAGSIA DFTGH 180
QY      121 APAAGIAPPCATVTHQADTSSPDPLFNPPLKTVGCOLDNANVTDAILBRAGSIA DFTGH 180
      121 APAAGIAPPCATVTHQADTSSPDPLFNPPLKTVGCOLDNANVTDAILBRAGSIA DFTGH 180
Db      181 YQTAPELELRLVNFPOSNCLKREKODESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
QY      181 YQTAPELELRLVNFPOSNCLKREKODESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
      181 YQTAPELELRLVNFPOSNCLKREKODESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
Db      241 EIFLLQAOAGMPGPGWGRITDISHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
QY      241 EIFLLQAOAGMPGPGWGRITDISHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
      241 EIFLLQAOAGMPGPGWGRITDISHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTA 300
Db      301 LTPHPPOKAYGVTLPTSVLFTAGHDNTLANLGALIELWNTLPGQPDNTPPGSELVFERM 360
QY      301 LTPHPPOKAYGVTLPTSVLFTAGHDNTLANLGALIELWNTLPGQPDNTPPGSELVFERM 360
      301 LTPHPPOKAYGVTLPTSVLFTAGHDNTLANLGALIELWNTLPGQPDNTPPGSELVFERM 360
Db      361 RLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGSEVKTLTAGCERNRQMGCSLAGFTQ 420
QY      361 RLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGSEVKTLTAGCERNRQMGCSLAGFTQ 420
      361 RLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGSEVKTLTAGCERNRQMGCSLAGFTQ 420
Db      421 IYNEARIPACSL 432
QY      421 IYNEARIPACSL 432
      421 IYNEARIPACSL 432
Db      421 IYNEARIPACSL 432

RESULT 13
US-11-018-709-1
; Sequence 1, Application US/11018709
; Publication No. US20050095691A1
; GENERAL INFORMATION:
; APPLICANT: Iel, Xingen
; TITLE OF INVENTION: SITE-DIRECTED MUTAGENESIS OF ESCHERICHIA COLI PHYTASE
; FILE REFERENCE: 19603/4031
; CURRENT APPLICATION NUMBER: US/11/018,709
; PRIOR FILING DATE: 2004-12-20
; PRIOR APPLICATION NUMBER: US/09/715,477
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,179
; PRIOR FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 432
; TYPE: prt
; ORGANISM: Escherichia coli
; US-11-018-709-1

Query Match      96.4%; Score 2177; DB 6; Length 432;
Best Local Similarity 97.9%; Pred. No. 4.3e-189;
Matches 423; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY      1 MKAIIIPFLSLIPLTPQSAFAQSEBELKESVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
      1 MKAIIIPFLSLIPLTPQSAFAQSEBELKESVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
Db      61 TWPVKLGELTPRGSELIAVIGHYWRQRLVADGLPKCGCPQSGQVAIADVDETRTKTGE 120
QY      61 TWPVKLGELTPRGSELIAVIGHYWRQRLVADGLPKCGCPQSGQVAIADVDETRTKTGE 120
      61 TWPVKLGELTPRGSELIAVIGHYWRQRLVADGLPKCGCPQSGQVAIADVDETRTKTGE 120
```

Db 61 TWPVGLGWLTRRGSELIAVLGHYQRRLVADGLLAKKCCPGCGVAAIADVERTKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOULNANVTDAIISRAGSIAIDFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOULNANVTDAIISRAGSIAIDFTGH 180
QY 181 YOTAFRELERLVANFPOSNCLCKREKODESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
Db 181 YOTAFRELERLVANFPOSNCLCKREKODESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
QY 241 EIFLLQOAGMPBEPMGRITDSHQMNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIKTA 300
Db 241 EIFLLQOAGMPBEPMGRITDSHQMNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIKTA 300
QY 301 LTPHPPOKQAVGVTLPSTVLFIAHDNTNLANLGGALBLNMTLPQOPDNTPPGSELVPERW 360
Db 301 LTPHPPOKQAVGVTLPSTVLFIAHDNTNLANLGGALBLNMTLPQOPDNTPPGSELVPERW 360
QY 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPPGEVKTLLAGCEERNAQMCSLAGFTQ 420
Db 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPPGEVKTLLAGCEERNAQMCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432

RESULT 14

US-09-777-566A-2
; Sequence 2, Application US/09777566A
; Patent No. US20010055788A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETTZ, Keith
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-6
; CURRENT APPLICATION NUMBER: US/09/777,566A
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1323)
; OTHER INFORMATION: n is any nucleotide
US-09-777-566A-2

Query Match 96.1%; Score 2171; DB 3; Length 440;
Best Local Similarity 97.7%; Pred. No. 1.6e-188;
Matches 422; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 KKAIIIPFLSLILPLTPGSAFAOSEPELKESVIVSRHGVRAPTKATQLMQDVTDPAMP 60
Db 1 KKAIIIPFLSLILPLTPGSAFAOSEPELKESVIVSRHGVRAPTKATQLMQDVTDPAMP 60
QY 61 TWPVGLGWLTRRGSELIAVLGHYQRRLVADGLLAKKCCPGCGVAAIADVERTKTGE 120
Db 61 TWPVGLGWLTRRGSELIAVLGHYQRRLVADGLLAKKCCPGCGVAAIADVERTKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOULNANVTDAIISRAGSIAIDFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOULNANVTDAIISRAGSIAIDFTGH 180

Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOULNANVTDAIISRAGSIAIDFTGH 180
QY 181 YOTAFRELERLVANFPOSNCLCKREKODESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
Db 181 YOTAFRELERLVANFPOSNCLCKREKODESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
QY 241 EIFLLQOAGMPBEPMGRITDSHQMNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIKTA 300
Db 241 EIFLLQOAGMPBEPMGRITDSHQMNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIKTA 300
QY 301 LTPHPPOKQAVGVTLPSTVLFIAHDNTNLANLGGALBLNMTLPQOPDNTPPGSELVPERW 360
Db 301 LTPHPPOKQAVGVTLPSTVLFIAHDNTNLANLGGALBLNMTLPQOPDNTPPGSELVPERW 360
QY 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPPGEVKTLLAGCEERNAQMCSLAGFTQ 420
Db 361 RRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPPGEVKTLLAGCEERNAQMCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432

RESULT 15

US-09-866-379-2
; Sequence 2, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETTZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-866-379-2

Query Match 96.1%; Score 2171; DB 3; Length 440;
Best Local Similarity 97.7%; Pred. No. 1.6e-188;
Matches 422; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 KKAIIIPFLSLILPLTPGSAFAOSEPELKESVIVSRHGVRAPTKATQLMQDVTDPAMP 60
Db 1 KKAIIIPFLSLILPLTPGSAFAOSEPELKESVIVSRHGVRAPTKATQLMQDVTDPAMP 60
QY 61 TWPVGLGWLTRRGSELIAVLGHYQRRLVADGLLAKKCCPGCGVAAIADVERTKTGE 120
Db 61 TWPVGLGWLTRRGSELIAVLGHYQRRLVADGLLAKKCCPGCGVAAIADVERTKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOULNANVTDAIISRAGSIAIDFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPPLKTGVCOULNANVTDAIISRAGSIAIDFTGH 180

```

Qy 181 YOTAPRELEERVLPQSNLCLKREKODESCSLTQALPSELKVSADCVSLTGAVSLASMLT 240
    |||||
Db 181 RQTAPELEERVLPQSNLCLKREKODESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
    |||||
Qy 241 EIFLQQAQGMPEPGMGRIITDSHQWMTLSLHNAQFDLQRTPEVARSRATPLLDLIXTA 300
    |||||
Db 241 EIFLQQAQGMPEPGMGRIITDSHQWMTLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
    |||||
Qy 301 LTPHPQOKQAYGVTLPTSVLPIAGHDTNLNLGALFLNMTLPQOPDNTPPGSELVPERW 360
    |||||
Db 301 LTPHPQOKQAYGVTLPTSVLPIAGHDTNLNLGALFLNMTLPQOPDNTPPGSELVPERW 360
    |||||
Qy 361 RRLSDNSQWIQVSLVFQTLQOMRDKTPLSLNTPPGSEVKLTLACGBERNNAQGMCSLAGFTQ 420
    |||||
Db 361 RRLSDNSQWIQVSLVFQTLQOMRDKTPLSLNTPPGSEVKLTLACGBERNNAQGMCSLAGFTQ 420
    |||||
Qy 421 IYNEARIPACSL 432
    |||||
Db 421 IYNEARIPACSL 432
    |||||

```

Search completed: June 13, 2006, 10:50:37
 Job time : 140.697 secs

THIS PAGE LEFT BLANK